



Diffusion approximation of critical controlled multi-type branching processes

Mátyás Barczy¹ · Miguel González² · Pedro Martín-Chávez² ·
Inés del Puerto²

Received: 18 April 2023 / Accepted: 19 March 2024 / Published online: 29 April 2024
© The Author(s) 2024

Abstract

Branching processes form an important family of stochastic processes that have been successfully applied in many fields. In this paper, we focus our attention on controlled multi-type branching processes (CMBPs). A Feller-type diffusion approximation is derived for some critical CMBPs. Namely, we consider a sequence of appropriately scaled random step functions formed from a critical CMBP with control distributions having expectations that satisfy a kind of linearity assumption. It is proved that such a sequence converges weakly toward a squared Bessel process supported by a ray determined by an eigenvector of a matrix related to the offspring mean matrix and the control distributions of the branching process in question. As applications, among others, we derive Feller-type diffusion approximations of critical, primitive multi-type branching processes with immigration and some two-sex branching processes. We also describe the asymptotic behaviour of the relative frequencies of distinct types of individuals for critical CMBPs.

Keywords Controlled branching processes · Multi-type branching processes · Two-sex branching processes · Diffusion approximation · Squared Bessel processes

Mathematics Subject Classification 60J80 · 60F17

✉ Pedro Martín-Chávez
pedromc@unex.es

Mátyás Barczy
barczy@math.u-szeged.hu

Miguel González
mvelasco@unex.es

Inés del Puerto
idelpuerto@unex.es

¹ HUN-REN–SZTE Analysis and Applications Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1., 6720 Szeged, Hungary

² Departamento de Matemáticas, Facultad de Ciencias, e Instituto de Computación Científica Avanzada (ICCAEx), Universidad de Extremadura, Av. de Elvas s/n, 06006 Badajoz, Spain

1 Introduction

Branching processes can be well-applied to describe evolutionary systems, where elements reproduce according to certain probability laws. These processes are commonly used in population dynamics, and, in this context, systems are referred to as populations and elements as individuals or cells. In this framework, multi-type branching processes, that are appropriate for modelling the evolution of populations in which different types of individuals coexist, have been successfully applied. For instance, these processes are used for modelling polymerase chain reaction (see, e.g., Sagitov and Ståhlberg [30]), for modelling cell proliferation kinetics (see, e.g., González et al. [15], Kimmel and Axelrod [26, Chapter 5] or Yanev [37]) or for studying extinction of outbreak of diseases (see, e.g., Mwasunda et al. [27]). We will focus our attention on the class of controlled multi-type branching processes (CMBPs). This class was first introduced in González et al. [14]. The key feature of CMBPs is that the number of parents of each type at a given generation is determined by a random control mechanism that depends on the number of individuals of different types in the previous generation.

CMBPs form a wide family of branching processes that include, as particular cases, the well-known classical branching processes, namely, multi-type branching processes without or with immigration (MBPs or MBPIs), or two-sex branching processes (see details in Example 2.1). Although, for all these processes, individuals give rise to offspring independently of each others, in general, a CMBP no longer satisfies the additive property (see, e.g., Athreya and Ney [4, Chapter 1, page 3]). This is an important property that is satisfied, for instance, by MBPs and MBPIs. The lack of the additive property for a general CMBP makes it difficult to study this kind of processes and it comes from the following fact. Provided that we know the number of different types of individuals in a generation, in case of a general CMBP, the number of different types of individuals in the next generation is a random sum of some independent random variables, while, in case of a MBP or a MBPI, it is a non-random sum (in the sense that the number of summands is deterministic).

This paper aims to obtain a Feller-type diffusion approximation for some critical CMBPs (details of the model are given in Sect. 2) and to study the asymptotic behaviour of the relative frequencies of distinct types of individuals. The problems under consideration have a dual motivation. On the one hand, it has interest in itself from a theoretical point of view. Our theoretical results (see Theorem 3.3 and its corollaries) may allow us to study relevant applied problems as well. For example, our result on the asymptotic behaviour of relative frequencies may find applications in cell kinetics. On the other hand, the obtained scaling limit theorem can enable us to carry out further research on statistical inference. For example, one may complete the study on weighted conditional least square estimators for some critical CMBPs, started in González et al. [11], or one might start to investigate more general quasi-likelihood estimators. This intended research could contribute to the modeling of real data using CMBPs.

From a mathematical point of view, achieving functional limit theorems for critical branching processes has attracted the interest of many researchers since 1951, when Feller [10] was able to provide the first formulation for Galton–Watson processes. He proved that the sequence of appropriately scaled critical Galton–Watson processes converges in distribution to a non-negative diffusion process without drift (for a detailed proof based on infinitesimal generators, see also Ethier and Kurtz [9, Theorem 9.1.3]). The extension to branching processes with immigration started with the pioneering paper of Wei and Winnicki [33], where the limit process is a squared Bessel process that can be characterized as the pathwise unique strong solution of a certain stochastic differential equation (SDE). Focusing on controlled branching models, a further extension to the critical, single-type case was first carried out by

Sriram et al. [32], and, more recently, by González et al. [13]. Furthermore, for some critical CMBPs, under quite involved technical conditions, González et al. [16, Corollary 4.1] proved a conditional weak limit theorem for the one-dimensional distributions provided that the explosion set has a positive probability. However, according to our knowledge, functional limit theorems for critical CMBPs are not available in the literature. Our present paper fills this gap, it is a natural extension of the result for critical single-type controlled branching processes (CBPs) in González et al. [13] to the multi-type case.

Assuming that the expectations of the control distributions satisfy a linear relationship with the population size additively perturbed by a function also depending on the population size (see (2.10)), we introduce a classification for such CMBPs based on the spectral radius of a matrix related to the offspring mean matrix and to the control distributions. Under some additional hypotheses, we prove that a suitably scaled and normalized critical CMBP converges weakly toward a squared Bessel process supported by a ray determined by an eigenvector of the aforementioned matrix, see Theorem 3.3. As corollaries, we are able to rediscover the known results on Feller-type diffusion approximations for critical, primitive MBPIs (see Corollaries 3.8 and 3.9). What is even more interesting is that we can apply our main theorem to get Feller-type diffusion approximations for some two-sex branching processes. We emphasize that no such results are available in the literature. Very recently, Bansaye et al. [5] have also proved a scaling limit theorem for a class of two-sex branching processes that combine classical asexual Galton–Watson processes and two-sex Galton–Watson branching processes introduced by Daley [8]. For a comparison of our results and theirs, see Remark 3.12. Finally, a result on the asymptotic behaviour of the relative frequencies for critical CMBPs is also derived from Theorem 3.3 (see Corollary 3.13). This kind of result has potential applications, for instance, in the field of cell kinetics, where it is more usual to measure relative frequencies instead of the absolute cell counts.

The proof of our main result (Theorem 3.3) follows the proof scheme of Theorem 3.1 in Ispány and Pap [21] for MBPIs, which is based on a weak convergence result for random step processes due to Ispány and Pap [20, Corollary 2.2] (see also Theorem A.4). This latter result has been applied in other papers to prove several scaling limit theorems, see, e.g., Ispány and Pap [21] and Ráth [28]. The lack of the additive property for a general CMBP makes the proof of our Theorem 3.3 more involved. Next, we outline the course of the proof, and we also point out the new ingredients in it. We start with determining the conditional moments of the branching process (see Proposition 2.3), which are essential to find out the asymptotic behaviour of some moments of the branching process (see Lemma A.3). The heart of the proof of Theorem 3.3 is an application of Theorem A.4 for a sequence of martingale differences formed from the CMBP in question (see (4.3) in Step 1). Then, in case of MBPIs, the Feller-type diffusion approximation follows straightforwardly from a continuous mapping theorem (see Theorem A.5), as Ispány and Pap [21] showed. However, this is not the case for general CMBPs, an extra additional work (see Step 3) is required. We also emphasize that the additive perturbation of the expectations of the control distributions mentioned above (see also (2.10)) results a new difficulty in the proofs compared to MBPIs, and it is addressed in the proofs of Steps 1 and 3.

The paper is structured as follows. In Sect. 2, CMBPs are defined, and, under the linearity assumption (2.11) on the conditional expectations, we introduce their classification by distinguishing subcritical, critical and supercritical CMBPs. By giving examples, we also point out that different types of classical branching processes can be viewed as particular cases of our model, thus illustrating its wide scope of applicability. In Sect. 3, we collect all the hypotheses that are assumed and we present all of our results obtained. Section 4 is devoted to the proof of Theorem 3.3, which is structured in four steps for an easy reading. We close

the paper with an Appendix which contains some auxiliary results such as the asymptotic behaviour of the first and second moments of CMBPs and the second and fourth moments of a corresponding martingale difference (see Lemma A.3).

2 Controlled multi-type branching processes

Let (Ω, \mathcal{F}, P) be a fixed probability space on which all the random variables will be defined, and let $\mathbb{N}, \mathbb{Z}_+, \mathbb{R}, \mathbb{R}_+$ and \mathbb{R}_{++} be the set of positive integers, non-negative integers, real numbers, non-negative real numbers, and positive real numbers, respectively. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, let us denote by $\mathbf{x} \leq \mathbf{y}$ if each coordinate of \mathbf{x} is less than or equal to the corresponding coordinate of \mathbf{y} . For $\mathbf{z} = (z_1, \dots, z_p)^\top \in \mathbb{R}^p$, let $|\mathbf{z}| := (|z_1|, \dots, |z_p|)^\top \in \mathbb{R}_+^p$, and $\mathbf{z}^+ := (z_1^+, \dots, z_p^+)^\top \in \mathbb{R}_+^p$, where x^+ stands for the positive part of $x \in \mathbb{R}$ and $^\top$ for the transpose. The natural basis in \mathbb{R}^p is denoted by $\mathbf{e}_1, \dots, \mathbf{e}_p$. The null vector in \mathbb{R}^p is denoted by $\mathbf{0}_p$, and $\mathbf{1}_p$ is for the vector in \mathbb{R}^p with all the coordinates 1. The Borel sigma-algebra on \mathbb{R}^p is denoted by $\mathcal{B}(\mathbb{R}^p)$. The trace of a matrix $A \in \mathbb{R}^{p \times p}$ is denoted by $\text{tr}(A)$. The $p \times p$ identity matrix is denoted by I_p . For a matrix $A \in \mathbb{R}^{l \times p}$, let $\text{Null}(A) := \{\mathbf{x} \in \mathbb{R}^p : A\mathbf{x} = \mathbf{0}_l\}$. Along the paper, we will not distinguish between the notations of the norm of a vector in \mathbb{R}^p and that of a matrix in $\mathbb{R}^{p \times p}$. For $\mathbf{z} \in \mathbb{R}^p$, let $\|\mathbf{z}\|$ denote the Euclidean norm of \mathbf{z} , and for a matrix $A \in \mathbb{R}^{p \times p}$, let $\|A\| := \max\{\|A\mathbf{z}\| : \|\mathbf{z}\| = 1, \mathbf{z} \in \mathbb{R}^p\}$. For a positive semi-definite matrix $A \in \mathbb{R}^{p \times p}$, let \sqrt{A} denote the unique symmetric positive semi-definite square root of A . For a function $\mathbf{h} : \mathbb{R}^p \rightarrow \mathbb{R}^p$, by the notation $\mathbf{h}(\mathbf{z}) = O(f(\mathbf{z}))$ as $\|\mathbf{z}\| \rightarrow \infty$, where $f : \mathbb{R}^p \rightarrow \mathbb{R}_+$, we mean that there exist $C > 0$ and $R > 0$ such that $\|\mathbf{h}(\mathbf{z})\| \leq Cf(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}^p$ with $\|\mathbf{z}\| > R$. Further, if for all $\epsilon > 0$ there exists $R > 0$ such that $\|\mathbf{h}(\mathbf{z})\| \leq \epsilon f(\mathbf{z})$ for all $\mathbf{z} \in \mathbb{R}^p$ with $\|\mathbf{z}\| > R$, then we write $\mathbf{h}(\mathbf{z}) = o(f(\mathbf{z}))$ as $\|\mathbf{z}\| \rightarrow \infty$. Convergence in probability is denoted by \xrightarrow{P} . Some further notations for weak convergence of stochastic processes with càdlàg sample paths are recalled in Sect. 3.

For a fixed $p \in \mathbb{N}$ and a \mathbb{Z}_+^p -valued random vector \mathbf{Z}_0 , let us consider a controlled p -type branching process $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$, defined recursively as

$$\mathbf{Z}_{k+1} := \sum_{i=1}^p \sum_{j=1}^p \phi_{k,i}(\mathbf{Z}_k) \mathbf{X}_{k,j,i}, \quad k \in \mathbb{Z}_+, \tag{2.1}$$

where $\mathbf{Z}_k, \phi_k(\mathbf{z}), \mathbf{z} \in \mathbb{Z}_+^p$, and $\mathbf{X}_{k,j,i}$ are \mathbb{Z}_+^p -valued random vectors:

$$\mathbf{Z}_k =: \begin{pmatrix} Z_{k,1} \\ \vdots \\ Z_{k,p} \end{pmatrix}, \quad \phi_k(\mathbf{z}) =: \begin{pmatrix} \phi_{k,1}(\mathbf{z}) \\ \vdots \\ \phi_{k,p}(\mathbf{z}) \end{pmatrix}, \quad \mathbf{X}_{k,j,i} =: \begin{pmatrix} X_{k,j,i,1} \\ \vdots \\ X_{k,j,i,p} \end{pmatrix}.$$

The intuitive interpretation of the process $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ is as follows:

- $Z_{k,i}$ is the number of i -type individuals in the k -th generation,
- $\phi_{k,i}(\mathbf{Z}_k)$ is the number of i -type progenitors in the k -th generation,
- $X_{k,j,i,l}$ is the number of l -type offsprings of the j -th i -type progenitor in the k -th generation.

Assume that $\{\mathbf{Z}_0, \phi_k(\mathbf{z}), \mathbf{X}_{k,j,i} : k \in \mathbb{Z}_+, j \in \mathbb{N}, \mathbf{z} \in \mathbb{Z}_+^p, i \in \{1, \dots, p\}\}$ are independent, $\{\phi_k(\mathbf{z}) : k \in \mathbb{Z}_+\}$ are identically distributed for each $\mathbf{z} \in \mathbb{Z}_+^p$ and $\{\mathbf{X}_{k,j,i} : k \in \mathbb{Z}_+, j \in \mathbb{N}\}$ are also identically distributed for each $i \in \{1, \dots, p\}$. The distributions of $\phi_0(\mathbf{z}), \mathbf{z} \in \mathbb{Z}_+^p$, are called the control distributions.

Moreover, for clarity of the readers, let us write out (2.1) for controlled two-type branching processes. The number of i -type individuals in the initial generation is $Z_{0,i}$, $i = 1, 2$, and the recursive definition in (2.1) takes the form

$$\begin{pmatrix} Z_{k+1,1} \\ Z_{k+1,2} \end{pmatrix} = \sum_{j=1}^{\phi_{k,1}(Z_{k,1}, Z_{k,2})} \begin{pmatrix} X_{k,j,1,1} \\ X_{k,j,1,2} \end{pmatrix} + \sum_{j=1}^{\phi_{k,2}(Z_{k,1}, Z_{k,2})} \begin{pmatrix} X_{k,j,2,1} \\ X_{k,j,2,2} \end{pmatrix}, \quad k \in \mathbb{Z}_+.$$

Part (iv) of Example 2.1 serves as a good illustration of a two-type case.

It is easy to check that the process defined in (2.1) is a \mathbb{Z}_+^p -valued Markov chain. This model is very general, several other popular branching processes can be seen as particular cases of the CMBPs, see Example 2.1.

Example 2.1 (i) Multi-type branching processes. We get the subclass of p -type branching processes without immigration by defining the deterministic control function as $\phi_k(z) := z$, $z \in \mathbb{Z}_+^p$, $k \in \mathbb{Z}_+$, in (2.1).

(ii) Multi-type branching processes with immigration. Let us consider such a MBPI $(Y_k)_{k \in \mathbb{Z}_+}$ given by

$$Y_{k+1} := \sum_{i=1}^p \sum_{j=1}^{Y_{k,i}} \xi_{k,j,i} + I_{k+1}, \quad k \in \mathbb{Z}_+, \tag{2.2}$$

where $\{Y_0, \xi_{k,j,i}, I_k : k \in \mathbb{Z}_+, j \in \mathbb{N}, i \in \{1, \dots, p\}\}$ are independent \mathbb{Z}_+^p -valued random vectors, $\{\xi_{k,j,i} : k \in \mathbb{Z}_+, j \in \mathbb{N}\}$ are identically distributed for each $i \in \{1, \dots, p\}$ (offspring distributions) and $\{I_k : k \in \mathbb{Z}_+\}$ are also identically distributed (immigration distribution). Note that $(Y_k)_{k \in \mathbb{Z}_+}$ can be written as a $(p+1)$ -type branching process with

$$Z_k := \begin{pmatrix} Y_k \\ 1 \end{pmatrix}, \quad X_{k,j,i} := \begin{pmatrix} \xi_{k,j,i} \\ 0 \end{pmatrix}, \quad X_{k,1,p+1} := \begin{pmatrix} I_{k+1} \\ 1 \end{pmatrix} \tag{2.3}$$

for $k \in \mathbb{Z}_+$, $j \in \mathbb{N}$, $i \in \{1, \dots, p\}$, when $\phi_k(z) := z$, $z \in \mathbb{Z}_+^{p+1}$, $k \in \mathbb{Z}_+$ (see (i)). Moreover, $(Y_k)_{k \in \mathbb{Z}_+}$ can also be written as a controlled $(p+1)$ -type branching process with the choices given in (2.3) and the control functions $\phi_k(z) = (z_1, \dots, z_p, 1)^\top$, $z \in \mathbb{Z}_+^{p+1}$, $k \in \mathbb{Z}_+$.

(iii) Multi-type branching processes with migration. Let us consider a stochastic process $(Z_k)_{k \in \mathbb{Z}_+}$ given by

$$Z_{k+1} := \sum_{i=1}^p \sum_{j=1}^{Z_{k,i} + M_{k,i}(Z_{k,i})} X_{k,j,i}, \quad k \in \mathbb{Z}_+,$$

where $\{Z_0, M_k(z), X_{k,j,i} : k \in \mathbb{Z}_+, j \in \mathbb{N}, z \in \mathbb{Z}_+^p, i \in \{1, \dots, p\}\}$ are independent \mathbb{Z}^p -valued random vectors such that Z_0 and $X_{k,j,i}$ have non-negative coordinates, $\{M_k(z) := (M_{k,1}(z_1), \dots, M_{k,p}(z_p))^\top : k \in \mathbb{Z}_+\}$ are identically distributed for each $z \in \mathbb{Z}_+^p$ with range contained in $[-z_1, \infty) \times \dots \times [-z_p, \infty)$, and $\{X_{k,j,i} : k \in \mathbb{Z}_+, j \in \mathbb{N}\}$ are also identically distributed for each $i \in \{1, \dots, p\}$. For each $i \in \{1, \dots, p\}$, $M_{k,i}(z_i)$ can be interpreted as a migration component for the i -type individuals in the k -th generation. Depending on the sign of $M_{k,i}(z_i)$, there is emigration (negative value), immigration (positive value) or no migration (zero value). In fact, this is an equivalent way of writing a CMBP, because every $\phi_k(z)$ can be written as $z + M_k(z)$, $z \in \mathbb{Z}_+^p$, $k \in \mathbb{N}$, with appropriate choices of $M_k(z)$.

- (iv) Two-sex Galton–Watson branching processes with immigration (2SBPIs). Let us consider such a process $(F_k, M_k)_{k \in \mathbb{Z}_+}$ with offspring and immigration distribution defined as follows. Let (F_0, M_0) be an \mathbb{Z}_+^2 -valued random variable (random initial generation) and

$$(F_{k+1}, M_{k+1}) := \sum_{j=1}^{U_k} (f_{k,j}, m_{k,j}) + (F_{k+1}^I, M_{k+1}^I), \quad U_k := L(F_k, M_k), \quad k \in \mathbb{Z}_+, \tag{2.4}$$

where $\{(F_0, M_0), (f_{k,j}, m_{k,j}), (F_k^I, M_k^I) : k \in \mathbb{Z}_+, j \in \mathbb{N}\}$ are independent \mathbb{Z}_+^2 -valued random vectors, $\{(f_{k,j}, m_{k,j}) : k \in \mathbb{Z}_+, j \in \mathbb{N}\}$ are identically distributed (offspring distribution), and $\{(F_k^I, M_k^I) : k \in \mathbb{Z}_+\}$ are also identically distributed (immigration distribution). Further, $(U_k)_{k \in \mathbb{Z}_+}$ is a sequence of mating units corresponding to the mating function $L : \mathbb{Z}_+ \times \mathbb{Z}_+ \rightarrow \mathbb{Z}_+$, assumed to be non-decreasing in each argument. In the terminology of Asmussen [3], L can be called a marriage function as well. Then the 2SBPI $(F_k, M_k)_{k \in \mathbb{Z}_+}$ can be considered a CMBP $(Z_k)_{k \in \mathbb{Z}_+}$ given by

$$Z_k := \begin{pmatrix} F_k \\ M_k \end{pmatrix}, \quad \phi_k(z) := \begin{pmatrix} L(z) \\ 1 \end{pmatrix}, \quad X_{k,j,1} := \begin{pmatrix} f_{k,j} \\ m_{k,j} \end{pmatrix}, \quad X_{k,j,2} := \begin{pmatrix} F_{k+1}^I \\ M_{k+1}^I \end{pmatrix}$$

for $k \in \mathbb{Z}_+, j \in \mathbb{N}$, and $z \in \mathbb{Z}_+^2$. In this case, we emphasize that the coordinates of the control can be interpreted in a different way compared to what is written after (2.1). Namely, the first coordinate of the control denotes the number of couples, while the second one denotes an extra couple, which brings in female and male immigrants. ■

Remark 2.2 As happens for the 2SBPI, the intuitive interpretation of the branching process that appears in part (ii) of Example 2.1 is lost when we see it as a $(p + 1)$ -type branching process or as a controlled $(p + 1)$ -type branching process. Indeed, we add the possibility that there is an extra type of individuals in the population, but the number of them in all generations is 1. The only one individual of extra type always gives birth one individual of extra type and possibly some individuals of other types, which correspond to the immigrants arriving in the population. Despite the fact that both cases lack practical interpretability, this mathematical description is perfectly valid and useful:

- (i) It will enable us to recover the known result on the asymptotic behaviour of critical, primitive MBPIs due to Ispány and Pap [21, Theorem 3.1] (see Corollary 3.8). Furthermore, if we do not want to lose the intuitive interpretation, we could consider the controlled p -type branching process $(Z_k)_{k \in \mathbb{Z}_+}$ with $Z_0 := Y_0, \phi_k(z) := z + I_{k+1}$, and $X_{k,j,i} := \xi_{k,j,i}$ for $k \in \mathbb{Z}_+, j \in \mathbb{N}$, and $z \in \mathbb{Z}_+^p$, where $\{Y_0, \xi_{k,j,i}, I_k : k \in \mathbb{Z}_+, j \in \mathbb{N}, i \in \{1, \dots, p\}\}$ are given in part (ii) of Example 2.1. This is the third representation of a MBPI (for the other two ones, see part (ii) of Example 2.1). Note also that $(Z_k)_{k \in \mathbb{Z}_+}$ is a particular case of the multi-type branching process with migration in part (iii) in Example 2.1. In the single-type critical case, González et al. [13, Remark 3.2.2] proved that (under some moment assumptions) $(n^{-1} Z_{\lfloor nt \rfloor})_{t \in \mathbb{R}_+}$ converges weakly as $n \rightarrow \infty$ toward a limit process, which coincides with the limit process for a critical (usual) single-type branching processes with immigration scaled and normalized in the same way. Corollary 3.9 will imply that a version of this statement remains true in the p -type critical case.
- (ii) It is the first time that a 2SBPI is written as a CMBP (see part (iv) of Example 2.1), such a rewriting has not been considered until now in the literature. In fact, since the

control distributions are deterministic in case of a 2SBPI, it suggests a possible way to generalize the notion of a 2SBPI by allowing random mating functions. Earlier, only two-sex Galton–Watson branching processes without immigration, introduced by Daley [8], were viewed as special controlled branching processes (see Sevast’yanov and Zubkov [31, model 3]). Furthermore, Theorem 3.3 together with part (iv) of Example 2.1 will allow us to obtain scaling limit theorems for some critical 2SBPIs previously not considered in the literature (see Corollary 3.10). ■

Let us introduce notations for some moments. In all what follows, we suppose that $E[\|X_{0,1,i}\|^4] < \infty$ for $i \in \{1, \dots, p\}$ and $E[\|\phi_0(z)\|^4] < \infty$ for $z \in \mathbb{Z}_+^p$, and we denote

$$m_i := E[X_{0,1,i}] \in \mathbb{R}_+^p, \quad \boldsymbol{\varepsilon}(z) := E[\phi_0(z)] \in \mathbb{R}_+^p, \quad (2.5)$$

$$\Sigma_i := \text{Var}[X_{0,1,i}] \in \mathbb{R}^{p \times p}, \quad \Gamma(z) := \text{Var}[\phi_0(z)] \in \mathbb{R}^{p \times p}, \quad (2.6)$$

$$\zeta_{i,l} := E[(X_{0,1,i,l} - m_{i,l})^4] \in \mathbb{R}_+, \quad \kappa_i(z) := E[(\phi_{0,i}(z) - \varepsilon_i(z))^4] \in \mathbb{R}_+, \quad (2.7)$$

where $i, l \in \{1, \dots, p\}$ and $z \in \mathbb{Z}_+^p$.

Note that the moments defined in (2.5) and (2.6) are of course well-defined under weaker assumptions, namely, under the existence of first and second order moments of the offspring and control distributions, respectively. We also remark that m_i , Σ_i and $\zeta_{i,l}$, $i, l \in \{1, \dots, p\}$, do not depend on the control distributions, while $\boldsymbol{\varepsilon}(z)$, $\Gamma(z)$, and $\kappa_i(z)$, $z \in \mathbb{Z}_+^p$, $i \in \{1, \dots, p\}$, do depend.

Let us consider the canonical filtration of the process $\mathcal{F}_k := \sigma(\mathbf{Z}_0, \dots, \mathbf{Z}_k)$, $k \in \mathbb{Z}_+$, introduce the matrix $\mathbf{m} := (m_1, \dots, m_p) \in \mathbb{R}_+^{p \times p}$, and the operator $\odot : \mathbb{R}^p \times (\mathbb{R}^{p \times p})^p \rightarrow \mathbb{R}^{p \times p}$, $\mathbf{z} \odot \mathbf{A} := \sum_{i=1}^p z_i \mathbf{A}_i$ for $\mathbf{z} = (z_1, \dots, z_p)^\top \in \mathbb{R}^p$ and $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_p) \in (\mathbb{R}^{p \times p})^p$.

The following proposition is a multi-type counterpart of Proposition 3.5 in González et al. [12].

Proposition 2.3 *For each $k \in \mathbb{N}$, we have*

$$E[\mathbf{Z}_k \mid \mathcal{F}_{k-1}] = \mathbf{m}\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1}), \quad (2.8)$$

$$\text{Var}[\mathbf{Z}_k \mid \mathcal{F}_{k-1}] = \boldsymbol{\varepsilon}(\mathbf{Z}_{k-1}) \odot \boldsymbol{\Sigma} + \mathbf{m}\Gamma(\mathbf{Z}_{k-1})\mathbf{m}^\top, \quad (2.9)$$

where $\boldsymbol{\Sigma} := (\Sigma_1, \dots, \Sigma_p) \in (\mathbb{R}^{p \times p})^p$.

Proof Let $k \in \mathbb{N}$ be fixed arbitrarily. By the Markov property, we get $E[\mathbf{Z}_k \mid \mathcal{F}_{k-1}] = E[\mathbf{Z}_k \mid \mathbf{Z}_{k-1}]$. Using that $\phi_{k-1}(z)$, $z \in \mathbb{Z}_+^p$, and $\mathbf{X}_{k-1,j,i}$, $j \in \mathbb{N}$, $i \in \{1, \dots, p\}$, are independent of each other and of \mathbf{Z}_{k-1} , we have

$$\begin{aligned} E[\mathbf{Z}_k \mid \mathbf{Z}_{k-1} = \mathbf{z}] &= \sum_{i=1}^p E \left[E \left[\sum_{j=1}^{\phi_{k-1,i}(\mathbf{z})} X_{k-1,j,i} \mid \phi_{k-1}(\mathbf{z}) \right] \right] \\ &= \sum_{i=1}^p E[\phi_{0,i}(z)] m_i = \mathbf{m}\boldsymbol{\varepsilon}(z) \end{aligned}$$

for $z \in \mathbb{Z}_+^p$, which implies (2.8). Now we turn to prove (2.9). Similarly as before, we get

$$\begin{aligned} & \text{Var} \left[\mathbf{Z}_k \mid \mathbf{Z}_{k-1} = \mathbf{z} \right] \\ &= \mathbb{E} \left[\text{Var} \left[\sum_{i=1}^p \sum_{j=1}^p \phi_{k-1,i}(\mathbf{z}) \mathbf{X}_{k-1,j,i} \mid \boldsymbol{\phi}_{k-1}(\mathbf{z}) \right] \right] + \text{Var} \left[\mathbb{E} \left[\sum_{i=1}^p \sum_{j=1}^p \phi_{k-1,i}(\mathbf{z}) \mathbf{X}_{k-1,j,i} \mid \boldsymbol{\phi}_{k-1}(\mathbf{z}) \right] \right] \\ &= \sum_{i=1}^p \mathbb{E}[\phi_{k-1,i}(\mathbf{z}) \Sigma_i] + \text{Var} \left[\sum_{i=1}^p \phi_{k-1,i}(\mathbf{z}) \mathbf{m}_i \right] = \boldsymbol{\varepsilon}(\mathbf{z}) \odot \boldsymbol{\Sigma} + \mathbf{m}\Gamma(\mathbf{z})\mathbf{m}^\top \end{aligned}$$

for $\mathbf{z} \in \mathbb{Z}_+^p$. This yields (2.9). □

From now on, we assume that there exist a matrix $\Lambda \in \mathbb{R}^{p \times p}$ and a function $\mathbf{h} : \mathbb{Z}_+^p \rightarrow \mathbb{R}^p$ with $\|\mathbf{h}(\mathbf{z})\| = o(\|\mathbf{z}\|)$ as $\|\mathbf{z}\| \rightarrow \infty$ such that

$$\boldsymbol{\varepsilon}(\mathbf{z}) = \Lambda \mathbf{z} + \mathbf{h}(\mathbf{z}), \quad \mathbf{z} \in \mathbb{Z}_+^p. \tag{2.10}$$

Roughly speaking, the assumption (2.10) means that the average quantity of ancestors can be expressed as a linear map of the number of individuals affected by a minor perturbation, which becomes insignificant (negligible) in comparison with the population size when this latter is large enough. It is worth noting that migration of parents, both emigration and immigration, is permitted under (2.10) (see, for instance, part (iii) of Example 2.1 assuming that $\|\mathbb{E}[\mathbf{M}_0(\mathbf{z})]\| = o(\|\mathbf{z}\|)$ as $\|\mathbf{z}\| \rightarrow \infty$). In the single-type case, i.e., in case of $p = 1$, a corresponding condition has already been assumed (see, e.g., the condition (i) in Theorems 1 and 2 in González et al. [17]). In the multi-type case, the assumption (2.10) was also considered in González and del Puerto [11, condition (11.2)] and in González et al. [16, condition (4.2)], but only with a diagonal matrix Λ having non-negative entries.

Using (2.8) and (2.10), we get

$$\mathbb{E} \left[\mathbf{Z}_k \mid \mathcal{F}_{k-1} \right] = \mathbf{m}\Lambda \mathbf{Z}_{k-1} + \mathbf{m}\mathbf{h}(\mathbf{Z}_{k-1}), \quad k \in \mathbb{N}. \tag{2.11}$$

We can introduce a classification for a subclass of CMBPs satisfying (2.11) based on the asymptotic behaviour of $\mathbb{E}[\mathbf{Z}_k]$ as $k \rightarrow \infty$. Taking expectations of both sides of (2.11), by a recursive argument, one can derive a formula for $\mathbb{E}[\mathbf{Z}_k]$, $k \in \mathbb{N}$ (see (A.5) in Appendix). From this, one can see that in the description of the asymptotic behaviour of $\mathbb{E}[\mathbf{Z}_k]$ as $k \rightarrow \infty$, the powers of the matrix $\tilde{\mathbf{m}} := \mathbf{m}\Lambda$ play a crucial role. We assume that $\tilde{\mathbf{m}} \in \mathbb{R}_+^{p \times p}$ and it has only one eigenvalue of maximum modulus $\tilde{\rho}$ with algebraic and geometric multiplicities equal one. By the spectral theory of matrices, the asymptotic behaviour of $(\mathbf{m}\Lambda)^k$ as $k \rightarrow \infty$ is determined by $\tilde{\rho}$, the other eigenvalues do not come into play. Motivated by this, by definition, we say that such a CMBP is subcritical if $\tilde{\rho} < 1$, critical if $\tilde{\rho} = 1$, and supercritical if $\tilde{\rho} > 1$. It is important to highlight that whenever the matrix $\tilde{\mathbf{m}}$ is primitive (i.e., there exists $n \in \mathbb{N}$ such that all the entries of $\tilde{\mathbf{m}}^n$ are positive), the Perron–Frobenius theorem (see also Lemma A.1) guarantees the unique existence of $\tilde{\rho}$. Moreover, the usual classification of multi-type branching processes without and with immigration coincides with the one proposed here (see, e.g., Athreya and Ney [4, Chapter V, page 186] and Kaplan [25, page 948]).

3 Results

Let $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ be a controlled p -type branching process given in (2.1). Let us introduce the following hypotheses.

Hypothesis 1 $E[\|Z_0\|^2]$, $E[\|X_{0,1,i}\|^4]$ and $E[\|\phi_0(z)\|^4]$ are finite for each $i \in \{1, \dots, p\}$ and $z \in \mathbb{Z}_+^p$.

Hypothesis 2 There exist $\Lambda \in \mathbb{R}^{p \times p}$, $\alpha \in \mathbb{R}^p$ and a function $g : \mathbb{Z}_+^p \rightarrow \mathbb{R}^p$ with $\|g(z)\| = o(1)$ as $\|z\| \rightarrow \infty$ such that

$$e(z) = \Lambda z + \alpha + g(z), \quad z \in \mathbb{Z}_+^p.$$

Hypothesis 3 $\|\Gamma(z)\| = o(\|z\|)$ as $\|z\| \rightarrow \infty$.

Hypothesis 4 $\kappa_i(z) = O(\|z\|^2)$ as $\|z\| \rightarrow \infty$ for $i = 1, \dots, p$.

Hypothesis 5 The matrix $\tilde{m} := m\Lambda$ belongs to $\mathbb{R}_+^{p \times p}$ and

- (a) $\tilde{\rho} := 1$ is an eigenvalue of \tilde{m} having algebraic and geometric multiplicities 1, and the absolute values of the other eigenvalues of \tilde{m} are less than 1,
- (b) there exist a unique right eigenvector $\tilde{u} \in \mathbb{R}_+^p$ and a unique left eigenvector $\tilde{v} \in \mathbb{R}_+^p$ corresponding to $\tilde{\rho} = 1$ such that $\tilde{u}_1 + \dots + \tilde{u}_p = 1$, $\Lambda \tilde{u} \in \mathbb{R}_+^p$ and $\tilde{v}^\top \tilde{u} = 1$,
- (c) $\lim_{k \rightarrow \infty} \tilde{m}^k = \tilde{\Pi}$ and there exist $\tilde{c} \in \mathbb{R}_{++}$ and $\tilde{r} \in (0, 1)$ such that $\|\tilde{m}^k - \tilde{\Pi}\| \leq \tilde{c}\tilde{r}^k$ for each $k \in \mathbb{N}$, where $\tilde{\Pi} := \tilde{u}\tilde{v}^\top$.

Hypothesis 6 For all $\epsilon > 0$ and $B > 0$, there exists $k_0(\epsilon, B) \in \mathbb{N}$ such that $P[\|Z_k\| \leq B] \leq \epsilon$ for each $k \geq k_0(\epsilon, B)$, $k \in \mathbb{N}$.

Remark 3.1 (i) In Hypotheses 1 and 4, the fourth order moment assumptions for the off-spring and control distributions are used in the proofs only for checking the conditional Lindeberg condition, namely, condition (iii) of Theorem A.4, in order to prove convergence of some random step processes toward a diffusion process. For critical single-type CBPs, assuming that the explosion set has probability one, a detailed exposition of a proof of the conditional Lindeberg condition in question under second order moment assumptions can be found in González et al. [13].

- (ii) If Hypothesis 2 holds for $\phi_0(z)$, $z \in \mathbb{Z}_+^p$, then it also holds for any of its linear transformations $B\phi_0(z) + \beta$, $z \in \mathbb{Z}_+^p$, with $B \in \mathbb{R}^{p \times p}$ and $\beta \in \mathbb{R}^p$, by replacing Λ , α and g with $B\Lambda$, $B\alpha + \beta$ and Bg , respectively. Note that, in Hypothesis 2, $\alpha \in \mathbb{R}^p$ can have negative coordinates as well, for an example, see part (i) of Remark 3.6.
- (iii) Taking into account Lemma A.1, we have that Hypothesis 5 holds if the matrix \tilde{m} is primitive, its Perron–Frobenius eigenvalue $\tilde{\rho}$ equals 1, and $\Lambda \tilde{u} \in \mathbb{R}_+^p$, where \tilde{u} is the right Perron–Frobenius eigenvector of \tilde{m} corresponding to $\tilde{\rho}$. Moreover, notice that if m is primitive, then a set of sufficient conditions for \tilde{m} to be primitive is that $\Lambda \in \mathbb{R}_+^{p \times p}$ and that m and Λ commute. Note also that if $\Lambda \in \mathbb{R}_+^{p \times p}$, then $\tilde{m} \in \mathbb{R}_+^{p \times p}$ and $\Lambda x \in \mathbb{R}_+^p$ for all $x \in \mathbb{R}_+^p$.
- (iv) A sufficient condition for Hypothesis 6 is the almost sure explosion of the process, i.e. $P[\|Z_k\| \rightarrow \infty \text{ as } k \rightarrow \infty] = 1$. Indeed, $\{\|Z_k\| \rightarrow \infty \text{ as } k \rightarrow \infty\} \subset \bigcup_{k=0}^\infty \bigcap_{n=k}^\infty \{\|Z_n\| > B\}$ for all $B > 0$. Therefore, by the continuity of probability, we have for all $B > 0$,

$$1 = P\left[\bigcup_{k=0}^\infty \bigcap_{n=k}^\infty \{\|Z_n\| > B\}\right] = \lim_{k \rightarrow \infty} P\left[\bigcap_{n=k}^\infty \{\|Z_n\| > B\}\right] \leq \liminf_{k \rightarrow \infty} P[\|Z_k\| > B],$$

yielding that $\lim_{k \rightarrow \infty} P[\|Z_k\| > B] = 1$ for all $B > 0$, or equivalently $\lim_{k \rightarrow \infty} P[\|Z_k\| \leq B] = 0$ for all $B > 0$. Consequently, Hypothesis 6 holds. It is also interesting to notice that Hypothesis 6 can not be derived from Hypotheses 1–5, see part (ii) of Remark 3.6 for a nontrivial example, where Hypotheses 1–5 hold, but Hypothesis 6 does not.

- (v) In case of $\mathbf{g} \equiv \mathbf{0}_p$, Hypothesis 6 is not needed for our forthcoming main Theorem 3.3. In fact, in the proof of Theorem 3.3 it will be seen that Hypothesis 6 is only used for deriving (4.34) and (4.37), which are trivially satisfied for $\mathbf{g} \equiv \mathbf{0}_p$. ■

Next, we give an example for a controlled p -type branching process for which the Hypotheses 1–5 hold.

Example 3.2 Let us consider a controlled p -type branching process with $E[\|Z_0\|^2] < \infty$, $E[\|X_{0,1,i}\|^4] < \infty$, $i \in \{1, \dots, p\}$, and

$$\phi_{k,i}(z) = z_i + U_{k,i} \mathbb{1}_{\{z_i > 0\}}, \quad k \in \mathbb{Z}_+, \quad z \in \mathbb{Z}_+^p, \quad i \in \{1, \dots, p\},$$

where $\{U_{k,i} : k \in \mathbb{Z}_+, i \in \{1, \dots, p\}\}$ are independent and identically distributed \mathbb{Z}_+^p -valued random variables such that their common distribution is the uniform distribution on the set $\{-1, 0, 1\}$, and $\mathbb{1}_A$ denotes the indicator function of a set A . The interpretation is as follows. Whenever there are i -type individuals, there is a migratory component $U_{k,i}$ that allows the emigration/immigration of one i -type progenitor or there could be no migration with the same probabilities $\frac{1}{3}$. This is a particular case of model (iii) in Example 2.1. Hypothesis 1 is trivially fulfilled and Hypotheses 2, 3 and 4 hold with $\Lambda = I_p$, $\alpha = \mathbf{0}_p$ and $\mathbf{g} \equiv \mathbf{0}_p$, since, for each $z \in \mathbb{Z}_+^p$, we have

$$\Gamma(z) = \left(\frac{2}{3} \delta_{i,j} \mathbb{1}_{\{z_i > 0, z_j > 0\}}\right)_{i,j=1}^p \quad \text{and} \quad \kappa_i(z) = \frac{2}{3} \mathbb{1}_{\{z_i > 0\}},$$

yielding that $\|\Gamma(z)\| \leq \frac{2}{3}$ and $\kappa_i(z) \leq \frac{2}{3}$ for $z \in \mathbb{Z}_+^p$. Moreover, taking into account parts (iii) and (iv) of Remark 3.1, if m is primitive with Perron–Frobenius eigenvalue 1, then Hypothesis 5 is also satisfied, and Hypothesis 6 is not needed. ■

Let $\mathbf{C}(\mathbb{R}_+, \mathbb{R}^p)$ be the space of \mathbb{R}^p -valued continuous functions on \mathbb{R}_+ and let us denote by $\mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$ the space of \mathbb{R}^p -valued càdlàg functions on \mathbb{R}_+ and by $\mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^p)$ its Borel σ -algebra corresponding to the metric defined in equation (1.26) in Jacod and Shiryaev [22, Chapter VI]. For \mathbb{R}^p -valued stochastic processes with càdlàg paths $(\mathcal{Y}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, and $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$, if the distribution of $(\mathcal{Y}_t^{(n)})_{t \in \mathbb{R}_+}$ on the space $(\mathbf{D}(\mathbb{R}_+, \mathbb{R}^p), \mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^p))$ converges weakly as $n \rightarrow \infty$ to the distribution of $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ on the same space, then we use the notation $(\mathcal{Y}_t^{(n)})_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ as $n \rightarrow \infty$.

Let us define a sequence of random step processes $(\mathcal{Z}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, using the controlled p -type branching process $(Z_k)_{k \in \mathbb{Z}_+}$, as

$$\mathcal{Z}_t^{(n)} := n^{-1} Z_{\lfloor nt \rfloor}, \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Recall that, under Hypothesis 5, $\tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \mathbb{R}_+^p$ are the right and left eigenvectors of \tilde{m} corresponding to eigenvalue $\tilde{\rho} = 1$, respectively.

Theorem 3.3 *Suppose that Hypotheses 1–6 hold for the CMBP $(Z_k)_{k \in \mathbb{Z}_+}$ given in (2.1). Then*

$$(\mathcal{Z}_t^{(n)})_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{Z}_t \tilde{\mathbf{u}})_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty, \tag{3.1}$$

where $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{Z}_t = \tilde{\mathbf{v}}^\top m \alpha dt + \sqrt{\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\mathbf{u}}) \odot \Sigma) \tilde{\mathbf{v}} \mathcal{Z}_t^+} dW_t, \quad t \in \mathbb{R}_+, \tag{3.2}$$

with initial value $\mathcal{Z}_0 = 0$, where $(W_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.

Remark 3.4 The coefficient functions $\mathbb{R} \ni x \mapsto \tilde{\mathbf{v}}^\top \mathbf{m}\boldsymbol{\alpha} \in \mathbb{R}$ and $\mathbb{R} \ni x \mapsto \sqrt{\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\boldsymbol{\mu}}) \odot \boldsymbol{\Sigma}) \tilde{\mathbf{v}} x^+} \in \mathbb{R}$ satisfy the conditions of Theorem 1 in Yamada and Watanabe [36], so pathwise uniqueness is guaranteed for SDE (3.2). Indeed, $\int_0^\epsilon u^{-1} du = \infty$ for all $\epsilon > 0$ and square root is a strictly increasing non-negative function vanishing at zero such that $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$ for all $x, y \in \mathbb{R}_+$. Moreover, if $\tilde{\mathbf{v}}^\top \mathbf{m}\boldsymbol{\alpha} \in \mathbb{R}_+$ (respectively, $\tilde{\mathbf{v}}^\top \mathbf{m}\boldsymbol{\alpha} > 0$), then \mathcal{Z}_t is non-negative (respectively, positive) for all $t > 0$ by the comparison theorem (see, e.g., Revuz and Yor [29, Chapter IX, Theorem 3.7]) and, consequently, \mathcal{Z}_t^+ in (3.2) may be replaced by \mathcal{Z}_t . Indeed, in case of $\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\boldsymbol{\mu}}) \odot \boldsymbol{\Sigma}) \tilde{\mathbf{v}} = 0$, we have $\mathcal{Z}_t = \tilde{\mathbf{v}}^\top \mathbf{m}\boldsymbol{\alpha} t$, $t \in \mathbb{R}_+$, and in case of $\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\boldsymbol{\mu}}) \odot \boldsymbol{\Sigma}) \tilde{\mathbf{v}} > 0$ and $t > 0$, we have \mathcal{Z}_t has a Gamma distribution with parameters $2\tilde{\mathbf{v}}^\top \mathbf{m}\boldsymbol{\alpha} / (\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\boldsymbol{\mu}}) \odot \boldsymbol{\Sigma}) \tilde{\mathbf{v}})$ and $2 / (\tilde{\mathbf{v}}^\top ((\Lambda \tilde{\boldsymbol{\mu}}) \odot \boldsymbol{\Sigma}) \tilde{\mathbf{v}} t)$ (following, e.g., from Ikeda and Watanabe [19, Chapter IV, Example 8.2]). ■

Remark 3.5 (i) It turns out that the limit process in (3.1) can also be characterized as the pathwise unique strong solution of the SDE

$$d\mathcal{Z}_t = \tilde{\Pi} \mathbf{m}\boldsymbol{\alpha} dt + \tilde{\Pi} \sqrt{(\Lambda \mathcal{Z}_t)^+ \odot \boldsymbol{\Sigma}} d\mathcal{W}_t, \quad t \in \mathbb{R}_+, \tag{3.3}$$

with initial value $\mathcal{Z}_0 = \mathbf{0}_p$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard p -dimensional Wiener process (for details, see Step 4 in the proof of Theorem 3.3).

(ii) If all the coordinates of the vector $\boldsymbol{\alpha} \in \mathbb{R}^p$ in Hypothesis 2 are negative, then it is easy to prove that the pathwise unique strong solution of the SDE (3.2) is $\mathcal{Z}_t = 0, t \in \mathbb{R}_+$. ■

In the next remark, we point out that Hypothesis 6 basically excludes that $\mathbb{P}[\boldsymbol{\phi}_0(\mathbf{0}_p) = \mathbf{0}_p] = 1$, this condition appears for the conditional weak limit theorems for one-dimensional distributions of critical CMBPs in González et al. [16, Sect. 4]. Further, we also demonstrate that Hypothesis 6 is independent of Hypotheses 1–5.

Remark 3.6 (i) If $\mathbb{P}[\boldsymbol{\phi}_0(\mathbf{0}_p) = \mathbf{0}_p] = 1$, then $\boldsymbol{\alpha} = -\mathbf{g}(\mathbf{0}_p)$ in Hypothesis 2. Moreover, if $\mathbb{P}[\boldsymbol{\phi}_0(\mathbf{0}_p) = \mathbf{0}_p] = 1$ and $\mathbf{0}_p$ is a state of $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$, i.e., there exists an $n_0 \in \mathbb{Z}_+$ such that $\mathbb{P}[\mathbf{Z}_{n_0} = \mathbf{0}_p] > 0$, then Hypothesis 6 does not hold. Indeed, in this case, we have $\{\mathbf{Z}_{n_0} = \mathbf{0}_p\} \subseteq \{\mathbf{Z}_k = \mathbf{0}_p\}$ for each $k \geq n_0, k \in \mathbb{Z}_+$, yielding that $\{\mathbf{Z}_{n_0} = \mathbf{0}_p\} \subseteq \{\|\mathbf{Z}_k\| \leq B\}$ for all $B > 0$ and $k \geq n_0, k \in \mathbb{Z}_+$. Consequently, we get $0 < \mathbb{P}[\mathbf{Z}_{n_0} = \mathbf{0}_p] \leq \mathbb{P}[\|\mathbf{Z}_k\| \leq B]$ for all $B > 0$ and $k \geq n_0, k \in \mathbb{Z}_+$. By choosing $\epsilon := \frac{1}{2} \mathbb{P}[\mathbf{Z}_{n_0} = \mathbf{0}_p] > 0$, one can see that Hypothesis 6 does not hold. If $\mathbb{P}[\boldsymbol{\phi}_0(\mathbf{0}_p) = \mathbf{0}_p] = 1$ and $\mathbf{0}_p$ is not a state of $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$, then, in principle, Hypothesis 6 can hold. For example, let $p := 2$,

$$\boldsymbol{\phi}_k(\mathbf{z}) =: \begin{cases} \mathbf{0}_2 & \text{if } \mathbf{z} = \mathbf{0}_2, \\ \begin{pmatrix} z_1 + 1 \\ z_1 + z_2 - 1 \end{pmatrix} & \text{if } \mathbf{z} \neq \mathbf{0}_2, \mathbf{z} \in \mathbb{Z}_+^2, \end{cases} \quad k \in \mathbb{Z}_+,$$

$$\mathbf{X}_{k,j,1} =: \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{X}_{k,j,2} =: \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{Z}_0 =: \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k \in \mathbb{Z}_+, j \in \mathbb{N}.$$

Then $\mathbb{P}[\boldsymbol{\phi}_0(\mathbf{0}_2) = \mathbf{0}_2] = 1$ and

$$\mathbf{Z}_k = (k + 1) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k \in \mathbb{N}.$$

Hence one can easily see that Hypotheses 1–6 hold with

$$\Lambda := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \alpha := \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{g}(z) := \begin{cases} \begin{pmatrix} -1 \\ 1 \end{pmatrix} & \text{if } z = \mathbf{0}_2, \\ \mathbf{0}_2 & \text{if } z \neq \mathbf{0}_2, z \in \mathbb{Z}_+^2, \end{cases}$$

and

$$\mathbf{m} = \tilde{\mathbf{m}} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{\mathbf{u}} := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{v}} := \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In this case, without the application of Theorem 3.3, we readily have that

$$(\mathcal{Z}_t^{(n)})_{t \in \mathbb{R}_+} = \left(\frac{1}{n} (\lfloor nt \rfloor + 1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} \left(t \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty.$$

Of course, Theorem 3.3 also gives the same result, since in this case the SDE (3.2) takes the form $d\mathcal{Z}_t = dt, t \in \mathbb{R}_+$, with initial value $\mathcal{Z}_0 = 0$, yielding that $\mathcal{Z}_t = t, t \in \mathbb{R}_+$.

- (ii) We give a nontrivial example, where Hypotheses 1–5 hold, but Hypothesis 6 does not. Fix $p := 2$ and $\mathbf{Z}_0 := \mathbf{1}_2$. For $k \in \mathbb{Z}_+$, let $\phi_k(z) := z + \mathbf{1}_2$ for $z \in \mathbb{Z}_+^2 \setminus \{\mathbf{0}_2\}$ and $\phi_k(\mathbf{0}_2) := \mathbf{0}_2$; and for $k, j \in \mathbb{N}$ and $i \in \{1, 2\}$, let $X_{k,j,i}$ be a random vector with values in $\{\mathbf{0}_2, \mathbf{1}_2\}$ taking both values with probability $\frac{1}{2}$. Then $\mathbf{m}_i = E[X_{0,1,i}] = \frac{1}{2} \mathbf{1}_2$ for $i \in \{1, 2\}$. Further, let $\Lambda := \mathbf{I}_2, \alpha := \mathbf{1}_2, \mathbf{g}(\mathbf{0}_2) := -\mathbf{1}_2$ and $\mathbf{g}(z) := \mathbf{0}_2$ for $z \in \mathbb{Z}_+^2 \setminus \{\mathbf{0}_2\}$. Hence $\tilde{\mathbf{m}} = \mathbf{m}\Lambda = \mathbf{m} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, which is a primitive matrix with Perron–Frobenius eigenvalue $\tilde{\rho} = 1$. Consequently, Hypotheses 1–5 hold with the above introduced $\Lambda, \alpha, \mathbf{g}, \tilde{\mathbf{m}}$, and $\tilde{\mathbf{u}} := \frac{1}{\sqrt{2}} \mathbf{1}_2 =: \tilde{\mathbf{v}}$. Moreover, since $\mathbf{0}_2$ is a state of $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ and $P[\phi_0(\mathbf{0}_2) = \mathbf{0}_2] = 1$, part (i) of this remark implies that Hypothesis 6 does not hold. ■

In the next remark, we point out that Theorem 3.3 with $p = 1$ gives back the result of González et al. [13] for critical single-type CBPs.

Remark 3.7 Theorem 3.3 with $p = 1$, i.e. for critical single-type CBPs, yields that $(\mathcal{Z}_t^{(n)})_{t \in \mathbb{R}_+}$ converges weakly toward $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ as $n \rightarrow \infty$, where $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{Z}_t = m\alpha dt + \sqrt{m^{-1}\Sigma\mathcal{Z}_t^+} d\mathcal{W}_t, \quad t \in \mathbb{R}_+,$$

with initial value $\mathcal{Z}_0 = 0$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process, and $m > 0$ and Σ are the offspring mean and variance, respectively (Λ must be m^{-1} in order that $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ be critical). It is worth mentioning that González et al. [13, Theorem 3.1] obtained the same limit process $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ assuming the following three hypotheses, namely, (C0): the explosion set has probability one, (C1): $\varepsilon(z) = m^{-1}z + m^{-1}\alpha' + o(1)$ as $z \rightarrow \infty$ with some $\alpha' > 0$, and (C2): $\Gamma(z) = o(z)$ as $z \rightarrow \infty$. Notice that if (C0), (C1) and (C2) hold, then our Hypotheses 2, 3, 5 and 6 are satisfied. Indeed, Hypotheses 2, 3 and 5 follow trivially from (C1) and (C2), and (C0) implies Hypothesis 6 as it was pointed out in part (iv) in Remark 3.1. Concerning Hypotheses 1 and 4, in part (i) of Remark 3.1, we already mentioned the reason why the fourth moments on the offspring and control distributions are assumed in our paper. ■

As a consequence of Theorem 3.3 one can deduce the asymptotic behaviour of a critical primitive MBPI previously proved by Ispány and Pap [21, Theorem 3.1]. Let $(\mathbf{Y}_k)_{k \in \mathbb{Z}_+}$ be a MBPI defined in (2.2) such that the offspring mean matrix $\mathbf{m}_\xi :=$

$(E[\xi_{0,1,1}], \dots, E[\xi_{0,1,p}]) \in \mathbb{R}_+^{p \times p}$ is primitive with Perron–Frobenius eigenvalue 1, and let \mathbf{u} and \mathbf{v} be its right and left Perron–Frobenius eigenvectors, respectively (see Lemma A.1). Then $(\mathbf{Y}_k)_{k \in \mathbb{Z}_+}$ is a critical primitive MBPI. Let us denote $\mathbf{m}_I := E[\mathbf{I}_1] \in \mathbb{R}_+^p$ and $\mathbf{V} := (\text{Var}[\xi_{0,1,1}], \dots, \text{Var}[\xi_{0,1,p}]) \in (\mathbb{R}^{p \times p})^p$.

Corollary 3.8 (Ispány and Pap [21, Theorem 3.1]) *Let $(\mathbf{Y}_k)_{k \in \mathbb{Z}_+}$ be a critical primitive p -type branching process with immigration such that $E[\|\mathbf{Y}_0\|^2] < \infty$, $E[\|\xi_{0,1,i}\|^4] < \infty$, $i \in \{1, \dots, p\}$, and $E[\|\mathbf{I}_1\|^4] < \infty$. Then*

$$(n^{-1}\mathbf{Y}_{[nt]})_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{Y}_t \mathbf{u})_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{Y}_t = \mathbf{v}^\top \mathbf{m}_I dt + \sqrt{\mathbf{v}^\top (\mathbf{u} \odot \mathbf{V}) \mathbf{v}} \mathcal{Y}_t^+ d\mathcal{W}_t, \quad t \in \mathbb{R}_+,$$

with initial value $\mathcal{Y}_0 = 0$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.

Proof To apply Theorem 3.3, let us rewrite $(\mathbf{Y}_k)_{k \in \mathbb{Z}_+}$ as a CMBP (see part (ii) of Example 2.1). For $k \in \mathbb{Z}_+$, $j \in \mathbb{N}$ and $i \in \{1, \dots, p\}$, we have

$$\mathbf{Z}_k = \begin{pmatrix} \mathbf{Y}_k \\ 1 \end{pmatrix}, \quad X_{k,j,i} = \begin{pmatrix} \xi_{k,j,i} \\ 0 \end{pmatrix}, \quad X_{k,1,p+1} = \begin{pmatrix} \mathbf{I}^{k+1} \\ 1 \end{pmatrix}.$$

Further, for $k \in \mathbb{Z}_+$ and $\mathbf{z} \in \mathbb{Z}_+^{p+1}$, the deterministic control functions $\phi_k(\mathbf{z}) = (z_1, \dots, z_p, 1)^\top$ can be written in the form $\phi_k(\mathbf{z}) = \Lambda \mathbf{z} + \alpha + \mathbf{g}(\mathbf{z})$, with

$$\Lambda = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_p \\ \mathbf{0}_p^\top & 0 \end{pmatrix}, \quad \alpha = \begin{pmatrix} \mathbf{0}_p \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{g} \equiv \mathbf{0}_{p+1}.$$

Then we have $E[\|\mathbf{Z}_0\|^2] = E[1 + \|\mathbf{Y}_0\|^2] < \infty$, $E[\|\mathbf{X}_{0,1,i}\|^4] = E[\|\xi_{0,1,i}\|^4] < \infty$, $i \in \{1, \dots, p\}$, $E[\|\mathbf{X}_{0,1,p+1}\|^4] = E[(1 + \|\mathbf{I}_1\|^2)^2] < \infty$. Thus $\Gamma(\mathbf{z}) = \mathbf{0} \in \mathbb{R}^{(p+1) \times (p+1)}$ and $\kappa_i(\mathbf{z}) = 0$, $\mathbf{z} \in \mathbb{Z}_+^{p+1}$, $i \in \{1, \dots, p\}$, and hence Hypotheses 1, 2, 3 and 4 hold. Since $\mathbf{g} \equiv \mathbf{0}_{p+1}$, Hypothesis 6 is not needed, see part (v) of Remark 3.1. Finally, Hypothesis 5 is satisfied with

$$\mathbf{m} = \begin{pmatrix} \mathbf{m}_\xi & \mathbf{m}_I \\ \mathbf{0}_p^\top & 1 \end{pmatrix}, \quad \tilde{\mathbf{m}} = \begin{pmatrix} \mathbf{m}_\xi & \mathbf{0}_p \\ \mathbf{0}_p^\top & 0 \end{pmatrix}, \quad \tilde{\mathbf{u}} = \begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{v}} = \begin{pmatrix} \mathbf{v} \\ 0 \end{pmatrix}, \quad \tilde{\Pi} = \begin{pmatrix} \Pi_\xi & \mathbf{0}_p \\ \mathbf{0}_p^\top & 0 \end{pmatrix},$$

where $\Pi_\xi := \lim_{k \rightarrow \infty} \mathbf{m}_\xi^k$, and \tilde{c} and \tilde{r} in part (c) of Hypothesis 5 can be chosen as for Π_ξ (since \mathbf{m}_ξ is primitive with Frobenius–Perron eigenvalue 1, and see part (iii) of Lemma A.1). Further, we get

$$\Sigma_i = \begin{pmatrix} \text{Var}[\xi_{0,1,i}] & \mathbf{0}_p \\ \mathbf{0}_p^\top & 0 \end{pmatrix}, \quad i \in \{1, \dots, p\}, \quad \Sigma_{p+1} = \begin{pmatrix} \text{Var}[\mathbf{I}_1] & \mathbf{0}_p \\ \mathbf{0}_p^\top & 0 \end{pmatrix},$$

and, consequently, Theorem 3.3 yields the statement. □

Motivated by part (i) of Remark 2.2, we derive another corollary of Theorem 3.3.

Corollary 3.9 *Let $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ be a critical controlled p -type branching process given in part (i) of Remark 2.2 such that $E[\|\mathbf{Y}_0\|^2] < \infty$, $E[\|\xi_{0,1,i}\|^4] < \infty$, $i \in \{1, \dots, p\}$, and $E[\|\mathbf{I}_1\|^4] < \infty$. Then*

$$(n^{-1}\mathbf{Z}_{[nt]})_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{Z}_t \mathbf{u})_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(Z_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$dZ_t = v^\top m_I dt + \sqrt{v^\top (u \odot V) v} Z_t^+ dW_t, \quad t \in \mathbb{R}_+,$$

with initial value $Z_0 = 0$, where $(W_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.

Proof We follow the same steps as in the proof of Corollary 3.8. We have $E[\|Z_0\|^2] = E[\|Y_0\|^2] < \infty$, $E[\|X_{0,1,i}\|^4] = E[\|\xi_{0,1,i}\|^4] < \infty$, $i \in \{1, \dots, p\}$, $E[\|\phi_0(z)\|^4] \leq 8 E[\|z\|^4 + \|I_1\|^4] < \infty$. Therefore, Hypothesis 1 is satisfied. Hypothesis 2 also holds because $e(z) = \Lambda z + \alpha + g(z)$, $z \in \mathbb{Z}_+^p$, with $\Lambda = I_p$, $\alpha = m_I$ and $g \equiv \mathbf{0}_p$. Due to $E[\|I_1\|^4] < \infty$, it is clear that $\Gamma(z) = \text{Var}[I_1] \in \mathbb{R}^{p \times p}$ and $\kappa_i(z) = E[(I_{1,i} - m_{I,i})^4] < \infty$, $z \in \mathbb{Z}_+^p$, $i \in \{1, \dots, p\}$, so Hypotheses 3 and 4 are satisfied. Furthermore, Hypothesis 5 holds with

$$m = m_\xi, \quad \tilde{m} = m_\xi, \quad \tilde{u} = u, \quad \tilde{v} = v, \quad \tilde{\Pi} = \Pi_\xi,$$

where $\Pi_\xi := \lim_{k \rightarrow \infty} m_\xi^k$, and \tilde{c} and \tilde{r} in part (c) of Hypothesis 5 can be chosen as for Π_ξ (see part (iii) of Lemma A.1). Indeed, by our assumption, m_ξ is primitive with Perron–Frobenius eigenvalue equals 1. Finally, Hypothesis 6 is not needed, since $g \equiv \mathbf{0}_p$ (see part (v) of Remark 3.1). Hence the result follows from Theorem 3.3 using also that $v^\top m_\xi m_I = v^\top m_I$, since v is the left Perron–Frobenius eigenvector for m_ξ . \square

Next, we establish a scaling limit theorem for a 2SBPI with the promiscuous mating function $L(x, y) := x \min\{1, y\}$, $x, y \in \mathbb{Z}_+$. This type of mating function was first considered by Daley [8], and it was deeply studied by Alsmeyer and Rösler [1, 2]. Up to our knowledge, no such result as the following one is available in the literature (for more details, see Remark 3.12).

Corollary 3.10 *Let $(F_k, M_k)_{k \in \mathbb{Z}_+}$ be a 2SBPI defined in (2.4) with the promiscuous mating function. Assume that (F_0, M_0) is a \mathbb{N}^2 -valued random vector, $E[\|(F_0, M_0)\|^2] < \infty$, $E[\|(f_{0,1}, m_{0,1})\|^4] < \infty$, and $E[\|(F_1^I, M_1^I)\|^4] < \infty$. If $P[M_1^I = 0] = 0$ and $E[f_{0,1}] = 1$, then*

$$(n^{-1}(F_{[nt]}, M_{[nt]}))_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{X}_t(1, E[m_{0,1}]))_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{X}_t = E[F_1^I] dt + \sqrt{\text{Var}[f_{0,1}]} \mathcal{X}_t^+ dW_t, \quad t \in \mathbb{R}_+,$$

with initial value $\mathcal{X}_0 = 0$, where $(W_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.

Proof Recall that $L(x, y) = x \min\{1, y\}$, $x, y \in \mathbb{Z}_+$.

First, using that $P[M_1^I = 0] = 0$ and (F_0, M_0) is \mathbb{N}^2 -valued, we check that $L(F_n, M_n) = F_n$ for each $n \in \mathbb{Z}_+$ almost surely. For this, since M_n is \mathbb{Z}_+ -valued, taking into account the form of L , it is enough to verify that $P[M_n = 0] = 0$, $n \in \mathbb{N}$. Indeed, by a conditioning argument with respect to U_n , we can get for each $n \in \mathbb{Z}_+$,

$$P[M_{n+1} = 0] = P[M_1^I = 0] E[(P[m_{0,1} = 0])^{U_n}] = 0.$$

Consequently, $(F_k, M_k)_{k \in \mathbb{Z}_+}$ coincides almost surely with the 2-type controlled branching process $(Z_k)_{k \in \mathbb{Z}_+}$ given by

$$Z_k := \begin{pmatrix} F_k \\ M_k \end{pmatrix}, \quad \phi_k(z) := \begin{pmatrix} z_1 \\ 1 \end{pmatrix}, \quad X_{k,j,1} := \begin{pmatrix} f_{k,j} \\ m_{k,j} \end{pmatrix}, \quad X_{k,j,2} := \begin{pmatrix} F_{k+1}^I \\ M_{k+1}^I \end{pmatrix}$$

for $k \in \mathbb{Z}_+$, $j \in \mathbb{N}$, and $\mathbf{z} = (z_1, z_2) \in \mathbb{Z}_+^2$.

We want to apply Theorem 3.3 for $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$. The control distributions are deterministic, so Hypothesis 3 and 4 are trivially satisfied. In addition, we get $E[\|\mathbf{Z}_0\|^2] = E[\|(F_0, M_0)\|^2] < \infty$, $E[\|\mathbf{X}_{0,1,1}\|^4] = E[\|(f_{0,1}, m_{0,1})\|^4] < \infty$ and $E[\|\mathbf{X}_{0,1,2}\|^4] = E[\|(F_1^I, M_1^I)\|^4] < \infty$, yielding that Hypothesis 1 holds. For each $\mathbf{z} \in \mathbb{Z}_+^2$, we have

$$\phi_0(\mathbf{z}) = \begin{pmatrix} z_1 \\ 1 \end{pmatrix} = \Lambda \mathbf{z} + \alpha \quad \text{with} \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and hence Hypothesis 2 holds with the given Λ , α , and $\mathbf{g} \equiv \mathbf{0}_2$. Hypothesis 6 is not needed, see part (v) of Remark 3.1. Further, we get

$$\begin{aligned} \mathbf{m} &= \begin{pmatrix} E[f_{0,1}] & E[F_1^I] \\ E[m_{0,1}] & E[M_1^I] \end{pmatrix}, & \Sigma_1 &= \begin{pmatrix} \text{Var}[f_{0,1}] & \text{Cov}[f_{0,1}, m_{0,1}] \\ \text{Cov}[f_{0,1}, m_{0,1}] & \text{Var}[m_{0,1}] \end{pmatrix}, \\ \tilde{\mathbf{m}} &= \begin{pmatrix} E[f_{0,1}] & 0 \\ E[m_{0,1}] & 0 \end{pmatrix}, & \Sigma_2 &= \begin{pmatrix} \text{Var}[F_1^I] & \text{Cov}[F_1^I, M_1^I] \\ \text{Cov}[F_1^I, M_1^I] & \text{Var}[M_1^I] \end{pmatrix}. \end{aligned}$$

Since $E[f_{0,1}] = 1$ and $\tilde{\mathbf{m}}^k = \tilde{\mathbf{m}}$, $k \in \mathbb{N}$, the conditions of Hypothesis 5 are satisfied with

$$\tilde{\mathbf{u}} = \frac{1}{1 + E[m_{0,1}]} \begin{pmatrix} 1 \\ E[m_{0,1}] \end{pmatrix}, \quad \tilde{\mathbf{v}} = \begin{pmatrix} 1 + E[m_{0,1}] \\ 0 \end{pmatrix}, \quad \Lambda \tilde{\mathbf{u}} = \frac{1}{1 + E[m_{0,1}]} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and $\tilde{\Pi} = \lim_{k \rightarrow \infty} \tilde{\mathbf{m}}^k = \tilde{\mathbf{m}}$. Therefore, Theorem 3.3 implies that

$$\left(\frac{1}{n} \begin{pmatrix} F_{[nt]} \\ M_{[nt]} \end{pmatrix} \right)_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} \left(\frac{\mathcal{Z}_t}{1 + E[m_{0,1}]} \begin{pmatrix} 1 \\ E[m_{0,1}] \end{pmatrix} \right)_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(\mathcal{Z}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{Z}_t = (1 + E[m_{0,1}]) E[F_1^I] dt + \sqrt{(1 + E[m_{0,1}]) \text{Var}[f_{0,1}]} \mathcal{Z}_t^+ d\mathcal{W}_t, \quad t \in \mathbb{R}_+,$$

with initial value $\mathcal{Z}_0 = 0$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process. With the process $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ given by $\mathcal{X}_t = (1 + E[m_{0,1}])^{-1} \mathcal{Z}_t$, $t \in \mathbb{R}_+$, we obtain the desired result. \square

Next, we apply Theorem 3.3 for a 2SBPI defined in (2.4) with a so-called self-fertilization mating function defined by $L(\mathbf{z}) := z_1 + z_2$, $\mathbf{z} = (z_1, z_2) \in \mathbb{R}_+^2$, which allows both females and males have parthenogenesis (asexual reproduction). In this model, every individual (regardless whether it is a female or male) corresponds to a mating unit and can have female and male offsprings independently of the other individuals. There is a kind of phenomenon in nature as well, e.g., for aphids and for some reptiles. The following result can be considered as another new contribution in the field of Feller-type diffusion approximations of 2SBPIs.

Corollary 3.11 *Let $(F_k, M_k)_{k \in \mathbb{Z}_+}$ be a 2SBPI defined in (2.4) with the self-fertilization mating function. Assume that $E[\|(F_0, M_0)\|^2] < \infty$, $E[\|(f_{0,1}, m_{0,1})\|^4] < \infty$, $E[\|(F_1^I, M_1^I)\|^4] < \infty$, and $E[f_{0,1}], E[m_{0,1}] \in (0, 1)$ are such that $E[f_{0,1}] + E[m_{0,1}] = 1$. Then*

$$(n^{-1}(F_{[nt]}, M_{[nt]}))_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{X}_t(E[f_{0,1}], E[m_{0,1}]))_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{X}_t = \left(E[F_1^I] + E[M_1^I] \right) dt + \sqrt{\text{Var}[f_{0,1} + m_{0,1}]} \mathcal{X}_t^+ d\mathcal{W}_t, \quad t \in \mathbb{R}_+,$$

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.

Proof Let us rewrite $(F_k, M_k)_{k \in \mathbb{Z}_+}$ as a CMBP $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ (see part (ii) of Example 2.1). We want to apply Theorem 3.3 for $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$. The control distributions are deterministic, so Hypothesis 3 and 4 are trivially satisfied. In addition, we get $E[\|\mathbf{Z}_0\|^2] = E[\|(F_0, M_0)\|^2] < \infty$, $E[\|\mathbf{X}_{0,1,1}\|^4] = E[\|(f_{0,1}, m_{0,1})\|^4] < \infty$ and $E[\|\mathbf{X}_{0,1,2}\|^4] = E[\|(F_1^I, M_1^I)\|^4] < \infty$, yielding that Hypothesis 1 holds. For each $\mathbf{z} \in \mathbb{Z}_+^2$, we have

$$\phi_0(\mathbf{z}) = \begin{pmatrix} z_1 + z_2 \\ 1 \end{pmatrix} = \Lambda \mathbf{z} + \boldsymbol{\alpha} \quad \text{with} \quad \Lambda = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\alpha} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and hence Hypothesis 2 holds with the given Λ , $\boldsymbol{\alpha}$, and $\mathbf{g} \equiv \mathbf{0}_2$. Hypothesis 6 is not needed, see part (v) of Remark 3.1. Further, we get

$$\begin{aligned} \mathbf{m} &= \begin{pmatrix} E[f_{0,1}] & E[F_1^I] \\ E[m_{0,1}] & E[M_1^I] \end{pmatrix}, & \Sigma_1 &= \begin{pmatrix} \text{Var}[f_{0,1}] & \text{Cov}[f_{0,1}, m_{0,1}] \\ \text{Cov}[f_{0,1}, m_{0,1}] & \text{Var}[m_{0,1}] \end{pmatrix}, \\ \tilde{\mathbf{m}} &= \begin{pmatrix} E[f_{0,1}] & E[f_{0,1}] \\ E[m_{0,1}] & E[m_{0,1}] \end{pmatrix}, & \Sigma_2 &= \begin{pmatrix} \text{Var}[F_1^I] & \text{Cov}[F_1^I, M_1^I] \\ \text{Cov}[F_1^I, M_1^I] & \text{Var}[M_1^I] \end{pmatrix}. \end{aligned}$$

The two eigenvalues of $\tilde{\mathbf{m}}$ are 0 and $E[f_{0,1}] + E[m_{0,1}] = 1$, and consequently the spectral radius of $\tilde{\mathbf{m}}$ is 1. Using also that $E[f_{0,1}], E[m_{0,1}] \in (0, 1)$, we have that $\tilde{\mathbf{m}}$ is a primitive matrix with Frobenius–Perron eigenvalue 1. Note also that $\tilde{\mathbf{m}}^k = \tilde{\mathbf{m}}$, $k \in \mathbb{N}$, since $E[f_{0,1}] + E[m_{0,1}] = 1$. Taking into these considerations, one can check that the conditions of Hypothesis 5 are satisfied with

$$\tilde{\mathbf{u}} = \begin{pmatrix} E[f_{0,1}] \\ E[m_{0,1}] \end{pmatrix}, \quad \tilde{\mathbf{v}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \Lambda \tilde{\mathbf{u}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \tilde{\Pi} = \lim_{k \rightarrow \infty} \tilde{\mathbf{m}}^k = \tilde{\mathbf{m}},$$

and the constants $\tilde{c} \in \mathbb{R}_{++}$ and $\tilde{r} \in (0, 1)$ in part (c) of Hypothesis 5 can be chosen arbitrarily. Therefore, Theorem 3.3 implies that

$$\left(\frac{1}{n} \begin{pmatrix} F_{[nt]} \\ M_{[nt]} \end{pmatrix} \right)_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} \left(\mathcal{X}_t \begin{pmatrix} E[f_{0,1}] \\ E[m_{0,1}] \end{pmatrix} \right)_{t \in \mathbb{R}_+} \quad \text{as } n \rightarrow \infty,$$

where $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$\begin{aligned} d\mathcal{X}_t &= \left(E[F_1^I] + E[M_1^I] \right) dt + \sqrt{(\text{Var}[f_{0,1}] + 2 \text{Cov}[f_{0,1}, m_{0,1}] + \text{Var}[m_{0,1}])} \mathcal{X}_t^+ d\mathcal{W}_t, \\ t &\in \mathbb{R}_+, \end{aligned}$$

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process. This easily yields the assertion. \square

Remark 3.12 Scaling limits for a class of processes that combine classical asexual Galton–Watson processes and two-sex Galton–Watson branching processes without immigration (introduced by Daley [8]) have been recently studied in Bansaye et al. [5]. However, this new family of processes does not include the class of two-sex Galton–Watson branching processes with immigration as a particular case, so the results in Bansaye et al. [5] do not imply our Corollaries 3.10 and 3.11. Even if we recover two-sex Galton–Watson branching processes without immigration as particular cases of the model (1)–(2) in Bansaye et al. [5]

(using their notations, with the special choices $\mathcal{E}_{n,p}^{f,N} := -1$ and $\mathcal{E}_{n,p}^{m,N} := -1$, $n, p \in \mathbb{N}$, which intuitively mean that the individuals (females or males) do not survive in the next generation), then their Theorem 2.1 cannot be applied, because their assumption (A1) does not hold in this case. Furthermore, note that for two-sex Galton–Watson branching processes without immigration, our Corollary 3.10 cannot be applied as well, since the assumption $P[M_1^f = 0] = 0$ does not hold in this case, however Corollary 3.11 can be applied and the limit process of the scaled two-sex Galton–Watson branching processes without immigration in question is the identically zero process. ■

Next, we formulate another corollary of Theorem 3.3, namely, we derive a limit distribution for the relative frequencies of distinct types of individuals. For different models, one can find similar results in Jagers [23, Corollary 1], in Yakovlev and Yanev [34, Proposition 1], [35, Theorem 2], and in Barczy and Pap [7, Corollary 4.1].

Recall that if \tilde{m} is primitive with Perron–Frobenius eigenvalue 1, then \tilde{u} and \tilde{v} denote the right and left Perron–Frobenius eigenvectors corresponding to 1, respectively.

Corollary 3.13 *Suppose that Hypotheses 1–4 and 6 hold for the CMBP $(Z_k)_{k \in \mathbb{Z}_+}$ given in (2.1). Assume also that the matrix \tilde{m} is primitive with Perron–Frobenius eigenvalue 1, and $\Lambda \tilde{u} \in \mathbb{R}_+^p$. If, in addition, $\tilde{v}^\top \alpha > 0$, then for all $t > 0$ and $i, j \in \{1, \dots, p\}$, we get*

$$\mathbb{1}_{\{e_j^\top Z_{[nt]} \neq 0\}} \frac{e_i^\top Z_{[nt]}}{e_j^\top Z_{[nt]}} \xrightarrow{P} \frac{e_i^\top \tilde{u}}{e_j^\top \tilde{u}} \quad \text{as } n \rightarrow \infty$$

and

$$\mathbb{1}_{\{Z_{[nt]} \neq \mathbf{0}_p\}} \frac{e_i^\top Z_{[nt]}}{\sum_{k=1}^p e_k^\top Z_{[nt]}} \xrightarrow{P} e_i^\top \tilde{u} \quad \text{as } n \rightarrow \infty.$$

Remark 3.14 The indicator functions $\mathbb{1}_{\{e_j^\top Z_{[nt]} \neq 0\}}$ and $\mathbb{1}_{\{Z_{[nt]} \neq \mathbf{0}_p\}}$ are needed in Corollary 3.13, since it can happen that $P[Z_{[nt]} = \mathbf{0}_p] > 0$ for some $t > 0$. ■

Proof of Corollary 3.13. For all $t > 0$ and $i, j \in \{1, \dots, p\}$, Theorem 3.3 yields that

$$\frac{1}{n} \left(e_i^\top Z_{[nt]}, e_j^\top Z_{[nt]} \right) \xrightarrow{\mathcal{L}} \left(e_i^\top \tilde{u} Z_t, e_j^\top \tilde{u} Z_t \right) \quad \text{as } n \rightarrow \infty.$$

The function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by $g(x, y) := \frac{x}{y} \mathbb{1}_{\{y \neq 0\}}$, $(x, y) \in \mathbb{R}^2$, is continuous on the set $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$, and the distribution of $(e_i^\top \tilde{u} Z_t, e_j^\top \tilde{u} Z_t)$ is concentrated on this set, since, by Remark 3.4, $P[Z_t > 0] = 1$ for all $t > 0$ and $e_j^\top \tilde{u} > 0$. Hence the continuous mapping theorem yields that

$$\begin{aligned} \mathbb{1}_{\{e_j^\top Z_{[nt]} \neq 0\}} \frac{e_i^\top Z_{[nt]}}{e_j^\top Z_{[nt]}} &= g(n^{-1} e_i^\top Z_{[nt]}, n^{-1} e_j^\top Z_{[nt]}) \\ &\xrightarrow{\mathcal{L}} g(e_i^\top \tilde{u} Z_t, e_j^\top \tilde{u} Z_t) = \mathbb{1}_{\{e_j^\top \tilde{u} Z_t \neq 0\}} \frac{e_i^\top \tilde{u} Z_t}{e_j^\top \tilde{u} Z_t} = \frac{e_i^\top \tilde{u}}{e_j^\top \tilde{u}} \quad \text{as } n \rightarrow \infty, \end{aligned}$$

thus we obtain the first convergence (using also that the limit is not random, so convergence in distribution yields convergence in probability in this case).

Similarly, for each $i \in \{1, \dots, p\}$, the function $h : \mathbb{R}^p \rightarrow \mathbb{R}$ defined by $h(x_1, \dots, x_p) := \frac{x_i}{\sum_{k=1}^p x_k} \mathbb{1}_{\{\sum_{k=1}^p x_k \neq 0\}}$, $(x_1, \dots, x_p) \in \mathbb{R}^p$, is continuous on the set $\mathbb{R}^p \setminus \{(x_1, \dots, x_p) \in \mathbb{R}^p : \sum_{k=1}^p x_k = 0\}$, and the distribution of $\tilde{u} Z_t$ is concentrated on this set, since $P[Z_t > 0] = 1$,

$t > 0$, and $\tilde{\mathbf{u}} \in \mathbb{R}_{++}^p$ imply that $\mathbb{P}[\sum_{k=1}^p \mathbf{e}_k^\top \tilde{\mathbf{u}} \mathcal{Z}_t = 0] = 0$. Using that $\mathbb{1}_{\{\mathbf{Z}_{\lfloor nt \rfloor} \neq \mathbf{0}_p\}} = \mathbb{1}_{\{\sum_{k=1}^p \mathbf{e}_k^\top \mathbf{Z}_{\lfloor nt \rfloor} \neq 0\}}$, the continuous mapping theorem yields that

$$\begin{aligned} \mathbb{1}_{\{\mathbf{Z}_{\lfloor nt \rfloor} \neq \mathbf{0}_p\}} \frac{\mathbf{e}_i^\top \mathbf{Z}_{\lfloor nt \rfloor}}{\sum_{k=1}^p \mathbf{e}_k^\top \mathbf{Z}_{\lfloor nt \rfloor}} &= h(n^{-1} \mathbf{Z}_{\lfloor nt \rfloor}) \\ \xrightarrow{\mathcal{L}} h(\tilde{\mathbf{u}} \mathcal{Z}_t) &= \mathbb{1}_{\{\sum_{k=1}^p \mathbf{e}_k^\top \tilde{\mathbf{u}} \mathcal{Z}_t \neq 0\}} \frac{\mathbf{e}_i^\top \tilde{\mathbf{u}} \mathcal{Z}_t}{\sum_{k=1}^p \mathbf{e}_k^\top \tilde{\mathbf{u}} \mathcal{Z}_t} = \frac{\mathbf{e}_i^\top \tilde{\mathbf{u}}}{\sum_{k=1}^p \mathbf{e}_k^\top \tilde{\mathbf{u}}} = \mathbf{e}_i^\top \tilde{\mathbf{u}} \end{aligned}$$

as $n \rightarrow \infty$, where we also used that the sum of coordinates of $\tilde{\mathbf{u}}$ is 1 (see part (ii) of Lemma A.1). Thus we obtain the second convergence as well (using again that the limit is not random). □

4 Proof of Theorem 3.3

The proof is divided in four steps. In the first three Steps 1, 2 and 3, we introduce some auxiliary stochastic processes, prove their convergence in distribution, and in Step 4, we discuss that how Steps 1–3 yield the assertion. The detailed proofs for Steps 1–3 can be found after Step 4.

Step 1

From the process $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$, we define a martingale difference sequence with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{Z}_+}$ given by

$$\mathbf{M}_k := \mathbf{Z}_k - \mathbb{E} \left[\mathbf{Z}_k \mid \mathcal{F}_{k-1} \right], \quad k \in \mathbb{N}, \tag{4.1}$$

and we also consider the following sequence of random step processes

$$\mathcal{M}_t^{(n)} = n^{-1} \left(\mathbf{Z}_0 + \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{M}_k \right), \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}. \tag{4.2}$$

We will prove that

$$\mathcal{M}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{M} \quad \text{as } n \rightarrow \infty, \tag{4.3}$$

where $(\mathcal{M}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{M}_t = \sqrt{(\Lambda \tilde{\Pi}(\mathcal{M}_t + t\mathbf{m}\alpha))^+} \odot \Sigma d\mathcal{W}_t, \quad t \in \mathbb{R}_+, \tag{4.4}$$

with initial value $\mathcal{M}_0 = \mathbf{0}_p$, where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a p -dimensional standard Wiener process.

Step 2

We will show that (4.3) implies

$$\Psi_n(\mathcal{M}^{(n)}) \xrightarrow{\mathcal{L}} \Psi(\mathcal{M}) \quad \text{as } n \rightarrow \infty, \tag{4.5}$$

where the mappings $\Psi : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$ and $\Psi_n : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $n \in \mathbb{N}$, are given by

$$\Psi(f)(t) := \tilde{\Pi}(f(t) + t\mathbf{m}\alpha), \tag{4.6}$$

$$\Psi_n(f)(t) := \tilde{m}^{\lfloor nt \rfloor} f(0) + \sum_{j=1}^{\lfloor nt \rfloor} \tilde{m}^{\lfloor nt \rfloor - j} \left(f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right) + \frac{1}{n}\mathbf{m}\alpha \right), \tag{4.7}$$

for $f \in \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$.

Step 3

We will check that $\mathcal{Z}^{(n)} = \Psi_n(\mathcal{M}^{(n)}) + \mathcal{V}^{(n)}$, $n \in \mathbb{N}$, and

$$\mathcal{Z}^{(n)} \xrightarrow{\mathcal{L}} \Psi(\mathcal{M}) \quad \text{as } n \rightarrow \infty, \tag{4.8}$$

where $(\mathcal{V}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, is another sequence of random step processes defined by

$$\mathcal{V}_t^{(n)} := n^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{m}^{\lfloor nt \rfloor - j} \mathbf{m}g(\mathbf{Z}_{j-1}), \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}. \tag{4.9}$$

Step 4

As a consequence of Steps 1–3, one can easily derive (3.1). Namely, let us define

$$\mathcal{Z}_t := \Psi(\mathcal{M}_t) = \tilde{\Pi}(\mathcal{M}_t + t\mathbf{m}\alpha), \quad t \in \mathbb{R}_+, \tag{4.10}$$

$$\mathcal{Y}_t := \tilde{\mathbf{v}}^\top \mathcal{Z}_t = \tilde{\mathbf{v}}^\top (\mathcal{M}_t + t\mathbf{m}\alpha), \quad t \in \mathbb{R}_+, \tag{4.11}$$

where we used that $\tilde{\mathbf{v}}^\top \tilde{\Pi} = \tilde{\mathbf{v}}^\top \tilde{\mathbf{u}} \tilde{\mathbf{v}}^\top = \tilde{\mathbf{v}}^\top$ (see Hypothesis 5).

Then $\mathcal{Z} = \mathcal{Y}\tilde{\mathbf{u}}$, since

$$\mathcal{Y}_t \tilde{\mathbf{u}} = \tilde{\mathbf{v}}^\top (\mathcal{M}_t + t\mathbf{m}\alpha) \tilde{\mathbf{u}} = \tilde{\mathbf{u}} (\tilde{\mathbf{v}}^\top \mathcal{M}_t) + t \tilde{\mathbf{u}} (\tilde{\mathbf{v}}^\top \mathbf{m}\alpha) = \tilde{\Pi} \mathcal{M}_t + t \tilde{\Pi} \mathbf{m}\alpha = \mathcal{Z}_t, \quad t \in \mathbb{R}_+.$$

Further, by Itô’s formula, we can show that $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE (3.2) with initial value 0, thus (3.1) follows and the proof concludes. Indeed, the SDE (4.4) can also be written in the form

$$d\mathcal{M}_t = \sqrt{(\tilde{\mathbf{v}}^\top (\mathcal{M}_t + t\mathbf{m}\alpha))^+ ((\Lambda \tilde{\mathbf{u}}) \odot \Sigma)} d\mathcal{W}_t, \quad t \in \mathbb{R}_+, \tag{4.12}$$

since $\Lambda \tilde{\mathbf{u}} \in \mathbb{R}_+^p$ (see Hypothesis 5) implies that

$$(\tilde{\mathbf{v}}^\top (\mathcal{M}_t + t\mathbf{m}\alpha))^+ \Lambda \tilde{\mathbf{u}} = ((\Lambda \tilde{\mathbf{u}}) \tilde{\mathbf{v}}^\top (\mathcal{M}_t + t\mathbf{m}\alpha))^+ = (\Lambda \tilde{\Pi} (\mathcal{M}_t + t\mathbf{m}\alpha))^+, \quad t \in \mathbb{R}_+.$$

Therefore, $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of the SDE

$$d\mathcal{Y}_t = \tilde{\mathbf{v}}^\top \mathbf{m}\alpha dt + \sqrt{\mathcal{Y}_t^+ \tilde{\mathbf{v}}^\top \sqrt{(\Lambda \tilde{\mathbf{u}}) \odot \Sigma}} d\mathcal{W}_t, \quad t \in \mathbb{R}_+, \tag{4.13}$$

with initial value 0, since, by (4.11), $d\mathcal{Y}_t = \tilde{\mathbf{v}}^\top \mathbf{m}\alpha dt + \tilde{\mathbf{v}}^\top d\mathcal{M}_t$, $t \in \mathbb{R}_+$, and $(\mathcal{M}_t)_{t \in \mathbb{R}_+}$ is the pathwise unique strong solution of (4.12) with initial value $\mathbf{0}_p$.

Suppose that $\tilde{\mathbf{v}}^\top((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)\tilde{\mathbf{v}} \neq 0$, then the process $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ given by

$$\mathcal{W}_t := \left(\tilde{\mathbf{v}}^\top((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)\tilde{\mathbf{v}}\right)^{-1/2} \tilde{\mathbf{v}}^\top \sqrt{(\Lambda\tilde{\mathbf{u}}) \odot \Sigma} \mathcal{W}_t, \quad t \in \mathbb{R}_+,$$

is a well-defined one-dimensional standard Wiener process, since $\Lambda\tilde{\mathbf{u}} \in \mathbb{R}_+^p$, yielding that $(\Lambda\tilde{\mathbf{u}}) \odot \Sigma$ is a positive semi-definite matrix. In such a case, the SDE (4.13) can be written as

$$d\mathcal{Y}_t = \tilde{\mathbf{v}}^\top \mathbf{m}\alpha \, dt + \sqrt{\mathcal{Y}_t^+} \sqrt{\tilde{\mathbf{v}}^\top((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)\tilde{\mathbf{v}}} d\mathcal{W}_t, \quad t \in \mathbb{R}_+, \tag{4.14}$$

which coincides with the SDE (3.2), as desired. Otherwise, it is trivial because if $\tilde{\mathbf{v}}^\top((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)\tilde{\mathbf{v}} = 0$, then $\|\tilde{\mathbf{v}}^\top \sqrt{(\Lambda\tilde{\mathbf{u}}) \odot \Sigma}\| = 0$, yielding that SDEs (3.2) and (4.13) correspond to the ODE $dx(t) = \tilde{\mathbf{v}}\mathbf{m}\alpha \, dt$, $t \in \mathbb{R}_+$, with initial value 0, and this ODE has the identically zero solution.

Finally, notice that the statement in Remark 3.5 can be derived using the equality $\mathcal{Z} = \mathcal{Y}\tilde{\mathbf{u}}$, (4.8), (4.10), (4.11), and (4.12). Namely, we have

$$d\mathcal{Z}_t = \tilde{\Pi}\mathbf{m}\alpha \, dt + \tilde{\Pi} d\mathcal{M}_t = \tilde{\Pi}\mathbf{m}\alpha \, dt + \tilde{\Pi} \sqrt{\mathcal{Y}_t^+((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)} d\mathcal{W}_t, \quad t \in \mathbb{R}_+.$$

Since $\Lambda\tilde{\mathbf{u}} \in \mathbb{R}_+^p$, we have

$$\mathcal{Y}_t^+ \Lambda\tilde{\mathbf{u}} = (\mathcal{Y}_t \Lambda\tilde{\mathbf{u}})^+ = (\Lambda\mathcal{Y}_t \tilde{\mathbf{u}})^+ = (\Lambda\mathcal{Z}_t)^+, \quad t \in \mathbb{R}_+,$$

yielding (3.3), as desired.

Proof of Step 1

We need to show that (4.3) holds. To this end, we apply Theorem A.4 with $\mathbf{b}(t, \mathbf{x}) = \mathbf{0}_p \in \mathbb{R}^p$, $\mathbf{C}(t, \mathbf{x}) = \sqrt{(\Lambda\tilde{\Pi}(\mathbf{x} + t\mathbf{m}\alpha))^+ \odot \Sigma} \in \mathbb{R}^{p \times p}$, $(t, \mathbf{x}) \in \mathbb{R}_+ \times \mathbb{R}^p$, $\boldsymbol{\eta}$ is the Dirac measure concentrated at $\mathbf{0}_p$, $\mathcal{U} = \mathcal{M}$, $\mathcal{F}_k^{(n)} = \mathcal{F}_k$, $\mathbf{U}_0^{(n)} = n^{-1}\mathbf{Z}_0$ and $\mathbf{U}_k^{(n)} = n^{-1}\mathbf{M}_k$, $k, n \in \mathbb{N}$.

First, we check that the SDE (4.4) has a pathwise unique strong solution for all \mathbb{R}^p -valued initial values. As seen in Step 4, we can rewrite the SDE (4.4) in the form (4.12).

If $(\mathcal{M}_t^{(y_0)})_{t \in \mathbb{R}_+}$ is a strong solution of (4.12) with initial value $\mathcal{M}_0^{(y_0)} = y_0 \in \mathbb{R}^p$, then we check that the process $((\mathcal{P}_t^{(y_0)}, \mathcal{Q}_t^{(y_0)}))_{t \in \mathbb{R}_+}$, defined by

$$\mathcal{P}_t^{(y_0)} := \tilde{\mathbf{v}}^\top(\mathcal{M}_t^{(y_0)} + t\mathbf{m}\alpha), \quad \mathcal{Q}_t^{(y_0)} := \mathcal{M}_t^{(y_0)} - \mathcal{P}_t^{(y_0)}\tilde{\mathbf{u}}, \quad t \in \mathbb{R}_+,$$

is a strong solution of the SDE

$$\begin{cases} d\mathcal{P}_t = \tilde{\mathbf{v}}^\top \mathbf{m}\alpha \, dt + \sqrt{\mathcal{P}_t^+} \tilde{\mathbf{v}}^\top \sqrt{(\Lambda\tilde{\mathbf{u}}) \odot \Sigma} d\mathcal{W}_t, \\ d\mathcal{Q}_t = -\tilde{\Pi}\mathbf{m}\alpha \, dt + \sqrt{\mathcal{P}_t^+} (I_p - \tilde{\Pi}) \sqrt{(\Lambda\tilde{\mathbf{u}}) \odot \Sigma} d\mathcal{W}_t, \end{cases} \quad t \in \mathbb{R}_+, \tag{4.15}$$

with initial value

$$(\mathcal{P}_0^{(y_0)}, \mathcal{Q}_0^{(y_0)}) = (\tilde{\mathbf{v}}^\top y_0, y_0 - \tilde{\mathbf{v}}^\top y_0 \tilde{\mathbf{u}}) = (\tilde{\mathbf{v}}^\top y_0, y_0 - \tilde{\mathbf{u}} \tilde{\mathbf{v}}^\top y_0) = (\tilde{\mathbf{v}}^\top y_0, (I_p - \tilde{\Pi})y_0).$$

Notice that the first SDE in (4.15) readily follows from (4.12), and the second one can be checked as follows

$$\begin{aligned} d\mathcal{Q}_t^{(y_0)} &= d\mathcal{M}_t^{(y_0)} - \tilde{\mathbf{u}} d\mathcal{P}_t^{(y_0)} = -\tilde{\mathbf{u}} \tilde{\mathbf{v}}^\top \mathbf{m}\alpha \, dt + (I_p - \tilde{\mathbf{u}} \tilde{\mathbf{v}}^\top) d\mathcal{M}_t^{(y_0)} \\ &= -\tilde{\Pi}\mathbf{m}\alpha \, dt + (I_p - \tilde{\Pi}) \sqrt{(\mathcal{P}_t^{(y_0)})^+((\Lambda\tilde{\mathbf{u}}) \odot \Sigma)} d\mathcal{W}_t, \quad t \in \mathbb{R}_+. \end{aligned}$$

Conversely, if $((\mathcal{P}_t^{(p_0, \mathbf{q}_0)}, \mathcal{Q}_t^{(p_0, \mathbf{q}_0)})_{t \in \mathbb{R}_+}$ is a strong solution of the SDE (4.15) with initial value $(\mathcal{P}_0^{(p_0, \mathbf{q}_0)}, \mathcal{Q}_0^{(p_0, \mathbf{q}_0)}) = (p_0, \mathbf{q}_0) \in \mathbb{R} \times \mathbb{R}^p$, again by Itô’s formula, the process

$$\mathcal{M}_t^{(p_0, \mathbf{q}_0)} := \mathcal{P}_t^{(p_0, \mathbf{q}_0)} \tilde{\mathbf{u}} + \mathcal{Q}_t^{(p_0, \mathbf{q}_0)}, \quad t \in \mathbb{R}_+,$$

is a strong solution of (4.12) with initial value $\mathcal{M}_0^{(p_0, \mathbf{q}_0)} = p_0 \tilde{\mathbf{u}} + \mathbf{q}_0$.

Now, let us see that the map $\mathbb{R}^p \ni \mathbf{y} \mapsto (\tilde{\mathbf{v}}^\top \mathbf{y}, (I_p - \tilde{\Pi}) \mathbf{y})$ is a bijection between \mathbb{R}^p and $\mathbb{R} \times \text{Null}(\tilde{\mathbf{v}}^\top)$. Let $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^p$ be two vectors with the same image under the map in question. Then the injectivity (i.e. $\mathbf{x}_0 = \mathbf{y}_0$) follows from

$$\begin{cases} \tilde{\mathbf{v}}^\top \mathbf{x}_0 = \tilde{\mathbf{v}}^\top \mathbf{y}_0 \\ (I_p - \tilde{\Pi}) \mathbf{x}_0 = (I_p - \tilde{\Pi}) \mathbf{y}_0, \end{cases} \quad \text{yielding that} \quad \begin{cases} \tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \mathbf{x}_0) = \tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \mathbf{y}_0) \\ \mathbf{x}_0 - \tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \mathbf{x}_0) = \mathbf{y}_0 - \tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \mathbf{y}_0), \end{cases}$$

where we have used Hypothesis 5. Further, for all $(p_0, \mathbf{q}_0) \in \mathbb{R} \times \text{Null}(\tilde{\mathbf{v}}^\top)$, the vector $\mathbf{y}_0 := p_0 \tilde{\mathbf{u}} + \mathbf{q}_0$ is an element of the preimage of (p_0, \mathbf{q}_0) , since, using that $\tilde{\mathbf{v}}^\top \tilde{\mathbf{u}} = 1$, we have $\mathbf{v}^\top (p_0 \tilde{\mathbf{u}} + \mathbf{q}_0) = p_0 + 0 = p_0$ and

$$(I_p - \tilde{\Pi})(p_0 \tilde{\mathbf{u}} + \mathbf{q}_0) = p_0 \tilde{\mathbf{u}} + \mathbf{q}_0 - p_0 \tilde{\mathbf{u}}(\tilde{\mathbf{v}}^\top \tilde{\mathbf{u}}) - \tilde{\mathbf{u}} \tilde{\mathbf{v}}^\top \mathbf{q}_0 = \mathbf{q}_0 - \tilde{\mathbf{u}} \cdot 0 = \mathbf{q}_0.$$

Hence the map in question is surjective.

All in all, it is enough to prove that SDE (4.15) has a pathwise unique strong solution for all initial values in $\mathbb{R} \times \mathbb{R}^p$, in particular for all initial values in $\mathbb{R} \times \text{Null}(\tilde{\mathbf{v}}^\top)$, which is shown below. The pathwise uniqueness for the first SDE in (4.15) (see also the SDE (4.13)) is clear by the discussion in Step 4 and Remark 3.4. We now proceed with the second SDE in (4.15). One can easily get that its pathwise unique strong solution with initial value $\mathcal{Q}_t^{(p_0, \mathbf{q}_0)} = \mathbf{q}_0$ takes the form

$$\mathcal{Q}_t^{(p_0, \mathbf{q}_0)} = \mathbf{q}_0 - \tilde{\Pi} \mathbf{m} \alpha t + (I_p - \tilde{\Pi}) \sqrt{(\Lambda \tilde{\mathbf{u}}) \odot \Sigma} \int_0^t \sqrt{(\mathcal{P}_s^{(p_0)})^+} d\mathcal{W}_s, \quad t \in \mathbb{R}_+,$$

and not only for all $(p_0, \mathbf{q}_0) \in \mathbb{R} \times \text{Null}(\tilde{\mathbf{v}}^\top)$ but also for all $(p_0, \mathbf{q}_0) \in \mathbb{R} \times \mathbb{R}^p$.

In what follows, we check that the assumptions of Theorem A.4 hold with our previous choices.

Step 1/A

For each $n, k \in \mathbb{N}$, we get that $\mathbb{E}[\|n^{-1} \mathbf{M}_k\|^2] < \infty$, since

$$\mathbb{E}[\|\mathbf{M}_k\|^2] = \text{tr}(\text{Var}[\mathbf{M}_k]) = \text{tr}\left(\mathbb{E}\left[\text{Var}\left[\mathbf{M}_k \mid \mathcal{F}_{k-1}\right]\right]\right) = \text{tr}\left(\mathbb{E}\left[\text{Var}\left[\mathbf{Z}_k \mid \mathcal{F}_{k-1}\right]\right]\right) < \infty,$$

where we used the variance decomposition formula and, at the last step, equation (2.9).

Further, $n^{-1} \mathbf{Z}_0 \xrightarrow{\mathcal{L}} \mathbf{0}_p$ as $n \rightarrow \infty$.

Step 1/B

For all $T > 0$, the condition (i) of Theorem A.4 is trivially satisfied, since $(\mathbf{M}_k)_{k \in \mathbb{N}}$ is a sequence of martingale differences yielding that $\mathbb{E}\left[\mathbf{M}_k \mid \mathcal{F}_{k-1}\right] = \mathbf{0}_p, k \in \mathbb{N}$.

The conditions (ii) and (iii) of Theorem A.4 can be written in the following forms

$$\sup_{t \in [0, T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left[\mathbf{M}_k (\mathbf{M}_k)^\top \mid \mathcal{F}_{k-1} \right] - \int_0^t (\mathcal{R}_s^{(n)})^+ ds \odot \Sigma \right\| \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow \infty, \tag{4.16}$$

$$\frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E} \left[\|\mathbf{M}_k\|^2 \mathbb{1}_{\{\|\mathbf{M}_k\| > n\theta\}} \mid \mathcal{F}_{k-1} \right] \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \rightarrow \infty \text{ for all } \theta > 0, \tag{4.17}$$

where the process $(\mathcal{R}_t^{(n)})_{t \in \mathbb{R}_+}$ is defined by

$$\mathcal{R}_t^{(n)} := \Lambda \tilde{\Pi} (\mathcal{M}_t^{(n)} + t\mathbf{m}\boldsymbol{\alpha}), \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Step 1/C

We show (4.16). Taking into account $\mathbb{E} \left[\mathbf{M}_k (\mathbf{M}_k)^\top \mid \mathcal{F}_{k-1} \right] = \text{Var} \left[\mathbf{Z}_k \mid \mathcal{F}_{k-1} \right], k \in \mathbb{N}$, equation (2.9) and Hypothesis 2, we can obtain

$$\begin{aligned} \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left[\mathbf{M}_k (\mathbf{M}_k)^\top \mid \mathcal{F}_{k-1} \right] &= \frac{1}{n^2} \left(\lfloor nt \rfloor \boldsymbol{\alpha} + \sum_{k=1}^{\lfloor nt \rfloor} (\Lambda \mathbf{Z}_{k-1} + \mathbf{g}(\mathbf{Z}_{k-1})) \right) \odot \Sigma \\ &\quad + \mathbf{m} \left(\frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \Gamma(\mathbf{Z}_{k-1}) \right) \mathbf{m}^\top, \quad t \in \mathbb{R}_+. \end{aligned} \tag{4.18}$$

Using again Hypothesis 2 together with equation (2.8), the martingale differences in (4.1) can be written as

$$\mathbf{M}_k = \mathbf{Z}_k - \tilde{\mathbf{m}} \mathbf{Z}_{k-1} - \mathbf{m}\boldsymbol{\alpha} - \mathbf{m}\mathbf{g}(\mathbf{Z}_{k-1}), \quad k \in \mathbb{N}. \tag{4.19}$$

With this expression and equation (4.2), we have

$$\begin{aligned} \mathcal{R}_t^{(n)} &= \Lambda \tilde{\Pi} \left(\frac{1}{n} \left(\mathbf{Z}_0 + \sum_{k=1}^{\lfloor nt \rfloor} (\mathbf{Z}_k - \tilde{\mathbf{m}} \mathbf{Z}_{k-1} - \mathbf{m}\boldsymbol{\alpha} - \mathbf{m}\mathbf{g}(\mathbf{Z}_{k-1})) \right) + t\mathbf{m}\boldsymbol{\alpha} \right) \\ &= \frac{1}{n} \Lambda \tilde{\Pi} \mathbf{Z}_{\lfloor nt \rfloor} + \frac{nt - \lfloor nt \rfloor}{n} \Lambda \tilde{\Pi} \mathbf{m}\boldsymbol{\alpha} - \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \Lambda \tilde{\Pi} \mathbf{m}\mathbf{g}(\mathbf{Z}_{k-1}), \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}, \end{aligned}$$

because $\tilde{\Pi} \tilde{\mathbf{m}} = (\lim_{l \rightarrow \infty} \tilde{\mathbf{m}}^l) \tilde{\mathbf{m}} = \lim_{l \rightarrow \infty} \tilde{\mathbf{m}}^{l+1} = \tilde{\Pi}$ by Hypothesis 5.

Since $\Lambda \tilde{\mathbf{u}} \in \mathbb{R}_+^p$, we have $\Lambda \tilde{\Pi} = (\Lambda \tilde{\mathbf{u}}) \tilde{\mathbf{v}}^\top \in \mathbb{R}_+^{p \times p}$ and, consequently, we get that

$$\mathcal{B}_{-,t}^{(n)} \leq (\mathcal{R}_t^{(n)})^+ \leq \mathcal{B}_{+,t}^{(n)}, \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_+,$$

where the bounds $\mathcal{B}_{-,t}^{(n)}$ and $\mathcal{B}_{+,t}^{(n)}$ are defined by

$$\mathcal{B}_{\pm,t}^{(n)} := \frac{1}{n} \Lambda \tilde{\Pi} \mathbf{Z}_{\lfloor nt \rfloor} \pm \frac{nt - \lfloor nt \rfloor}{n} |\Lambda \tilde{\Pi} \mathbf{m}\boldsymbol{\alpha}| \pm \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} |\Lambda \tilde{\Pi} \mathbf{m}\mathbf{g}(\mathbf{Z}_{k-1})|,$$

and recall that $|z| = (|z_1|, \dots, |z_p|)^\top \in \mathbb{R}_+^p$ for any $z = (z_1, \dots, z_p)^\top \in \mathbb{R}^p$. Hence we get

$$\int_0^t \mathcal{B}_{-,s}^{(n)} ds \leq \int_0^t (\mathcal{R}_s^{(n)})^+ ds \leq \int_0^t \mathcal{B}_{+,s}^{(n)} ds, \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}. \tag{4.20}$$

Next, we check that, for all $t \in \mathbb{R}_+$,

$$\begin{aligned} \int_0^t \mathcal{B}_{\pm,s}^{(n)} ds &= \frac{1}{n^2} \sum_{j=0}^{\lfloor nt \rfloor - 1} \Lambda \tilde{\Pi} \mathbf{Z}_j + \frac{nt - \lfloor nt \rfloor}{n^2} \Lambda \tilde{\Pi} \mathbf{Z}_{\lfloor nt \rfloor} \pm \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2}{2n^2} |\Lambda \tilde{\Pi} \mathbf{m} \boldsymbol{\alpha}| \\ &\pm \frac{1}{n^2} \sum_{j=0}^{\lfloor nt \rfloor - 1} \sum_{k=1}^j |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})| \pm \frac{nt - \lfloor nt \rfloor}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})|. \end{aligned} \tag{4.21}$$

To verify (4.21), we split the integral as follows

$$\int_0^t \mathcal{B}_{\pm,s}^{(n)} ds = \sum_{j=0}^{\lfloor nt \rfloor - 1} \int_{\frac{j}{n}}^{\frac{j+1}{n}} \mathcal{B}_{\pm,s}^{(n)} ds + \int_{\frac{\lfloor nt \rfloor}{n}}^t \mathcal{B}_{\pm,s}^{(n)} ds,$$

and we take into account

$$\int_{\frac{j}{n}}^{\frac{j+1}{n}} \mathcal{B}_{\pm,s}^{(n)} ds = \frac{1}{n^2} \Lambda \tilde{\Pi} \mathbf{Z}_j \pm \frac{1}{2n^2} |\Lambda \tilde{\Pi} \mathbf{m} \boldsymbol{\alpha}| \pm \frac{1}{n^2} \sum_{k=1}^j |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})|$$

for $j \in \{0, \dots, \lfloor nt \rfloor - 1\}$, and

$$\begin{aligned} \int_{\frac{\lfloor nt \rfloor}{n}}^t \mathcal{B}_{\pm,s}^{(n)} ds &= \frac{nt - \lfloor nt \rfloor}{n^2} \Lambda \tilde{\Pi} \mathbf{Z}_{\lfloor nt \rfloor} \pm \frac{(nt - \lfloor nt \rfloor)^2}{2n^2} |\Lambda \tilde{\Pi} \mathbf{m} \boldsymbol{\alpha}| \\ &\pm \frac{nt - \lfloor nt \rfloor}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})|. \end{aligned}$$

In what follows, we will use that for all $\mathbf{A} \in \mathbb{R}^{p \times p}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p, \mathbf{y}_0 \in \mathbb{R}_+^p$ with $\mathbf{y} - \mathbf{y}_0 \leq \mathbf{x} \leq \mathbf{y} + \mathbf{y}_0$, we have

$$\|\mathbf{A} - \mathbf{x} \odot \boldsymbol{\Sigma}\| \leq \|\mathbf{A} - \mathbf{y} \odot \boldsymbol{\Sigma}\| + \|\mathbf{y}_0\| \|\boldsymbol{\Sigma}\|, \tag{4.22}$$

where we recall $\|\boldsymbol{\Sigma}\| = \sum_{i=1}^p \|\boldsymbol{\Sigma}_i\|$. Hence, from equations (4.18), (4.20), (4.21) and (4.22), we obtain

$$\begin{aligned} &\left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E} \left[\mathbf{M}_k (\mathbf{M}_k)^\top \mid \mathcal{F}_{k-1} \right] - \int_0^t (\mathcal{R}_s^{(n)})^+ ds \odot \boldsymbol{\Sigma} \right\| \\ &\leq \left\| \frac{1}{n^2} \sum_{k=0}^{\lfloor nt \rfloor - 1} \Lambda (I_p - \tilde{\Pi}) \mathbf{Z}_k \odot \boldsymbol{\Sigma} \right\| + \left\| \frac{nt - \lfloor nt \rfloor}{n^2} \Lambda \tilde{\Pi} \mathbf{Z}_{\lfloor nt \rfloor} \odot \boldsymbol{\Sigma} \right\| \\ &+ \left\| \frac{\lfloor nt \rfloor}{n^2} \boldsymbol{\alpha} \odot \boldsymbol{\Sigma} \right\| + \left\| \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2}{2n^2} |\Lambda \tilde{\Pi} \mathbf{m} \boldsymbol{\alpha}| \right\| \|\boldsymbol{\Sigma}\| \\ &+ \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbf{g}(\mathbf{Z}_{k-1}) \odot \boldsymbol{\Sigma} \right\| + \left\| \mathbf{m} \left(\frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \Gamma(\mathbf{Z}_{k-1}) \right) \mathbf{m}^\top \right\| \\ &+ \left\| \frac{1}{n^2} \sum_{j=0}^{\lfloor nt \rfloor - 1} \sum_{k=1}^j |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})| \right\| \|\boldsymbol{\Sigma}\| + \left\| \frac{nt - \lfloor nt \rfloor}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} |\Lambda \tilde{\Pi} \mathbf{m} \mathbf{g}(\mathbf{Z}_{k-1})| \right\| \|\boldsymbol{\Sigma}\|. \end{aligned} \tag{4.23}$$

As a consequence of Hypothesis 2, the function \mathbf{g} is bounded, and $\| |A\mathbf{z}| \| = \|A\mathbf{z}\| \leq \|A\| \|\mathbf{z}\|$ for all $A \in \mathbb{R}^{p \times p}$ and $\mathbf{z} \in \mathbb{R}^p$. Thus, in order to show (4.16), we can easily see that it is enough to prove that

$$\sup_{t \in [0, T]} \frac{1}{n^2} \sum_{k=0}^{\lfloor nt \rfloor - 1} \|(l_p - \tilde{\Pi})\mathbf{Z}_k\| \xrightarrow{P} 0, \tag{4.24}$$

$$\sup_{t \in [0, T]} \frac{1}{n^2} \|\mathbf{Z}_{\lfloor nt \rfloor}\| \xrightarrow{P} 0, \tag{4.25}$$

$$\sup_{t \in [0, T]} \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \|\Gamma(\mathbf{Z}_{k-1})\| \xrightarrow{P} 0, \tag{4.26}$$

$$\sup_{t \in [0, T]} \frac{1}{n^2} \sum_{j=0}^{\lfloor nt \rfloor - 1} \sum_{k=1}^j \|\mathbf{g}(\mathbf{Z}_{k-1})\| \xrightarrow{P} 0, \tag{4.27}$$

as $n \rightarrow \infty$ for all $T > 0$. Indeed, for all $T > 0$, the supremum on $[0, T]$ of the fifth and eight terms on the right hand side of (4.23) tend to 0 as $n \rightarrow \infty$ in probability, since

$$\begin{aligned} \sup_{t \in [0, T]} \left\| \frac{nt - \lfloor nt \rfloor}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} |\mathbf{g}(\mathbf{Z}_{k-1})| \right\| &\leq \sup_{t \in [0, T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} |\mathbf{g}(\mathbf{Z}_{k-1})| \right\| \\ &\leq \frac{\lfloor nT \rfloor}{n^2} \sup_{z \in \mathbb{Z}_+^p} \|\mathbf{g}(z)\| = O\left(\frac{1}{n}\right) \end{aligned}$$

as $n \rightarrow \infty$. We remark that, later, at the end of the proof of Step 3, it will turn out that $n^{-1} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\mathbf{g}(\mathbf{Z}_{k-1})\|] \rightarrow 0$ as $n \rightarrow \infty$ for all $T > 0$ holds as well.

Let us start proving (4.24) and (4.25). From equation (4.19), the following expression can be obtained recursively,

$$\mathbf{Z}_k = \tilde{m}^k \mathbf{Z}_0 + \sum_{j=1}^k \tilde{m}^{k-j} (\mathbf{M}_j + \mathbf{m}\alpha + \mathbf{m}\mathbf{g}(\mathbf{Z}_{j-1})), \quad k \in \mathbb{N}, \tag{4.28}$$

and, using it with $\tilde{\Pi}\tilde{m}^k = \tilde{\Pi}$, $k \in \mathbb{Z}_+$, we have

$$(l_p - \tilde{\Pi})\mathbf{Z}_k = (\tilde{m}^k - \tilde{\Pi})\mathbf{Z}_0 + \sum_{j=1}^k (\tilde{m}^{k-j} - \tilde{\Pi}) (\mathbf{M}_j + \mathbf{m}\alpha + \mathbf{m}\mathbf{g}(\mathbf{Z}_{j-1})), \quad k \in \mathbb{N}.$$

Let $\tilde{c} \in \mathbb{R}_{++}$ and $\tilde{r} \in (0, 1)$ be the constants given in part (c) of Hypothesis 5 such that $\|\tilde{m}^k - \tilde{\Pi}\| \leq \tilde{c}\tilde{r}^k$ for each $k \in \mathbb{N}$, and let us introduce the constants

$$\tilde{C} := \sup_{j \in \mathbb{Z}_+} \|\tilde{m}^j\| \leq \tilde{c} + \|\tilde{\Pi}\|, \quad C_g := \sup_{z \in \mathbb{Z}_+^p} \|\mathbf{g}(z)\| < \infty. \tag{4.29}$$

Then, using (4.28), the following inequalities hold

$$\begin{aligned} \|\mathbf{Z}_{\lfloor nt \rfloor}\| &\leq \|\tilde{m}^{\lfloor nt \rfloor}\| \|\mathbf{Z}_0\| + \sum_{j=1}^{\lfloor nt \rfloor} \|\tilde{m}^{\lfloor nt \rfloor - j}\| \left(\|\mathbf{M}_j\| + \|\mathbf{m}\|(\|\alpha\| + \|\mathbf{g}(\mathbf{Z}_{j-1})\|) \right) \\ &\leq \tilde{C} \left(\|\mathbf{Z}_0\| + \lfloor nt \rfloor \|\mathbf{m}\|(\|\alpha\| + C_g) + \sum_{j=1}^{\lfloor nt \rfloor} \|\mathbf{M}_j\| \right), \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_+, \end{aligned}$$

and

$$\begin{aligned} \sum_{k=0}^{\lfloor nt \rfloor - 1} \|(l_p - \tilde{\Pi})\mathbf{Z}_k\| &\leq \tilde{c} \sum_{k=0}^{\lfloor nt \rfloor - 1} \tilde{r}^k \|\mathbf{Z}_0\| + \tilde{c} \sum_{k=0}^{\lfloor nt \rfloor - 1} \sum_{j=1}^k \tilde{r}^{k-j} \left(\|\mathbf{M}_j\| \right. \\ &\quad \left. + \|\mathbf{m}\| \times (\|\boldsymbol{\alpha}\| + \|\mathbf{g}(\mathbf{Z}_{j-1})\|) \right) \\ &\leq \frac{\tilde{c}}{1 - \tilde{r}} \left(\|\mathbf{Z}_0\| + \lfloor nt \rfloor \|\mathbf{m}\| (\|\boldsymbol{\alpha}\| + C_g) + \sum_{j=1}^{\lfloor nt \rfloor} \|\mathbf{M}_j\| \right), \end{aligned}$$

for all $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$, where at the last inequality we also used that

$$\sum_{k=0}^{\lfloor nt \rfloor - 1} \sum_{j=1}^k \tilde{r}^{k-j} x_j \leq \frac{1}{1 - \tilde{r}} \sum_{j=1}^{\lfloor nt \rfloor} x_j \quad \text{for all } \mathbf{x} = (x_1, \dots, x_{\lfloor nt \rfloor})^\top \in \mathbb{R}_+^{\lfloor nt \rfloor}.$$

Hence, to get (4.24) and (4.25) it is enough to show

$$\frac{1}{n^2} \|\mathbf{Z}_0\| \xrightarrow{P} 0, \quad \frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} \|\mathbf{M}_k\| \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty, \quad \text{for all } T > 0. \quad (4.30)$$

The first convergence in (4.30) holds trivially, since, in fact, $n^{-2} \|\mathbf{Z}_0\| \rightarrow 0$ as $n \rightarrow \infty$ almost surely.

The second convergence in (4.30) follows from (A.3), since, by Jensen inequality, $E[\|\mathbf{M}_k\|] \leq \sqrt{E[\|\mathbf{M}_k\|^2]} = O(k^{1/2})$ as $k \rightarrow \infty$, and hence, for all $T > 0$, we get that

$$\sum_{k=1}^{\lfloor nT \rfloor} E[\|\mathbf{M}_k\|] = O(n^{3/2}) \quad \text{as } n \rightarrow \infty.$$

Consequently, $n^{-2} \sum_{k=1}^{\lfloor nT \rfloor} \|\mathbf{M}_k\|$ converges to 0 in L^1 as $n \rightarrow \infty$.

Next, we check convergence (4.26). It is enough to prove that for all $T > 0$, we have that

$$\frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\Gamma(\mathbf{Z}_{k-1})\|] \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (4.31)$$

By Hypothesis 3, for all $\epsilon > 0$ there exists $K(\epsilon) > 0$ such that $\|\Gamma(\mathbf{z})\| < \epsilon \|\mathbf{z}\|$ for each $\mathbf{z} \in \mathbb{Z}_+^p$ with $\|\mathbf{z}\| > K(\epsilon)$. Since

$$E[\|\Gamma(\mathbf{Z}_{k-1})\|] = E[\|\Gamma(\mathbf{Z}_{k-1})\| (\mathbb{1}_{\{\|\mathbf{Z}_{k-1}\| \leq K(\epsilon)\}} + \mathbb{1}_{\{\|\mathbf{Z}_{k-1}\| > K(\epsilon)\}})], \quad k \in \mathbb{N},$$

for all $T > 0, \epsilon > 0$ and $n \in \mathbb{N}$, we have

$$\begin{aligned} \frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\Gamma(\mathbf{Z}_{k-1})\|] &\leq \frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} \left(\max_{\{\mathbf{z} \in \mathbb{Z}_+^p : \|\mathbf{z}\| \leq K(\epsilon)\}} \|\Gamma(\mathbf{z})\| + \epsilon E[\|\mathbf{Z}_{k-1}\|] \right) \\ &\leq \frac{T}{n} \max_{\{\mathbf{z} \in \mathbb{Z}_+^p : \|\mathbf{z}\| \leq K(\epsilon)\}} \|\Gamma(\mathbf{z})\| + \frac{\epsilon}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\mathbf{Z}_{k-1}\|], \end{aligned}$$

where $\max_{\{z \in \mathbb{Z}_+^p : \|z\| \leq K(\epsilon)\}} \|\Gamma(z)\| < \infty$. Using part (i) of Lemma A.3, we get $\frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\mathbf{Z}_{k-1}\|] = O(1)$ as $n \rightarrow \infty$. Hence for all $T > 0$ and $\epsilon > 0$, we get

$$\lim_{n \rightarrow \infty} \left(\frac{T}{n} \max_{\{z \in \mathbb{Z}_+^p : \|z\| \leq K(\epsilon)\}} \|\Gamma(z)\| + \frac{\epsilon}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} E[\|\mathbf{Z}_{k-1}\|] \right) = 0,$$

which yields (4.31), as desired.

Similarly, we prove (4.27) by checking that for all $T > 0$, we get that

$$\frac{1}{n^2} \sum_{j=1}^{\lfloor nT \rfloor} \sum_{k=0}^{j-1} E[\|\mathbf{g}(\mathbf{Z}_k)\|] \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{4.32}$$

By Hypothesis 2, for all $\epsilon > 0$ there exists $N(\epsilon) > 0$ such that $\|\mathbf{g}(z)\| < \epsilon$ for all $z \in \mathbb{Z}_+^p$ with $\|z\| > N(\epsilon)$. Then, using the notation C_g introduced in (4.29), we get

$$\begin{aligned} E[\|\mathbf{g}(\mathbf{Z}_k)\|] &= E[\|\mathbf{g}(\mathbf{Z}_k)\| (\mathbf{1}_{\{\|\mathbf{Z}_k\| \leq N(\epsilon)\}} + \mathbf{1}_{\{\|\mathbf{Z}_k\| > N(\epsilon)\}})] \\ &\leq C_g P[\|\mathbf{Z}_k\| \leq N(\epsilon)] + \epsilon, \quad k \in \mathbb{N}, \end{aligned} \tag{4.33}$$

and then using Hypothesis 6 we can derive that for all $T > 0$ and $\epsilon > 0$,

$$\begin{aligned} \sum_{j=1}^{\lfloor nT \rfloor} \sum_{k=0}^{j-1} E[\|\mathbf{g}(\mathbf{Z}_k)\|] &\leq \epsilon \lfloor nT \rfloor^2 + C_g \sum_{k=0}^{\lfloor nT \rfloor - 1} (\lfloor nT \rfloor - k) P[\|\mathbf{Z}_k\| \leq N(\epsilon)] \\ &\leq \epsilon \lfloor nT \rfloor^2 + C_g \sum_{k=0}^{k_0(\epsilon, N(\epsilon)) - 1} (\lfloor nT \rfloor - k) P[\|\mathbf{Z}_k\| \leq N(\epsilon)] \\ &\quad + C_g \sum_{k=k_0(\epsilon, N(\epsilon))}^{\lfloor nT \rfloor - 1} (\lfloor nT \rfloor - k) \epsilon \end{aligned} \tag{4.34}$$

for sufficiently large $n \in \mathbb{N}$, where

$$\sum_{k=0}^{k_0(\epsilon, N(\epsilon)) - 1} (\lfloor nT \rfloor - k) P[\|\mathbf{Z}_k\| \leq N(\epsilon)] \leq k_0(\epsilon, N(\epsilon)) \lfloor nT \rfloor$$

and

$$\sum_{k=k_0(\epsilon, N(\epsilon))}^{\lfloor nT \rfloor - 1} (\lfloor nT \rfloor - k) \epsilon \leq \epsilon n^2 T^2.$$

As a consequence, for all $T > 0$ and $\epsilon > 0$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{j=1}^{\lfloor nT \rfloor} \sum_{k=0}^{j-1} E[\|\mathbf{g}(\mathbf{Z}_k)\|] \leq \epsilon T^2 + \epsilon C_g T^2.$$

Hence, by taking the limit as $\epsilon \downarrow 0$, we obtain (4.32), as desired.

Step 1/D

We show (4.17). For this, it is enough to verify that for all $T > 0$ and $\theta > 0$, we have

$$\frac{1}{n^2} \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}[\|M_k\|^2 \mathbb{1}_{\{\|M_k\| > n\theta\}}] \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{4.35}$$

Using that

$$\mathbb{E}[\|M_k\|^2 \mathbb{1}_{\{\|M_k\| > n\theta\}}] \leq \frac{1}{n^2 \theta^2} \mathbb{E}[\|M_k\|^4], \quad n, k \in \mathbb{N},$$

the convergence (4.35) follows from equation (A.4). Consequently, we finished the proof of (4.3).

Proof of Step 2

We want to apply Theorem A.5 to prove the convergence (4.5). We need to check that the assumptions of Theorem A.5 are satisfied. The continuity of Ψ can be straightforwardly deduced (following also from Jacod and Shiryaev [22, Chapter VI, Proposition 1.23]), so the measurability of Ψ holds. For the sequence $(\Psi_n)_{n \in \mathbb{N}}$ and $N \in \mathbb{N}$, let us introduce the localized sequence $(\Psi_n^N)_{n \in \mathbb{N}}$ given by $\Psi_n^N : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $\Psi_n^N(f)(t) := \Psi_n(f)(t \wedge N)$ for $f \in \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$. Since, for each $n \in \mathbb{N}$ and $f \in \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $\Psi_n^N(f) \rightarrow \Psi_n(f)$ as $N \rightarrow \infty$, it is enough to check the measurability of Ψ_n^N , $n \in \mathbb{N}$ (see Barczy et al. [6, page 603] for the details). Briefly, for each $n \in \mathbb{N}$, one can introduce the auxiliary measurable mappings $\Psi_n^{N,1} : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow (\mathbb{R}^p)^{nN+1}$ and $\Psi_n^{N,2} : (\mathbb{R}^p)^{nN+1} \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$ defined by

$$\Psi_n^{N,1}(f) := \left(f(0), f\left(\frac{1}{n}\right), f\left(\frac{2}{n}\right), \dots, f(N) \right)$$

for $f \in \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, and

$$\Psi_n^{N,2}(x_0, x_1, \dots, x_{nN})(t) := \tilde{m}^{\lfloor n(t \wedge N) \rfloor} x_0 + \sum_{j=1}^{\lfloor n(t \wedge N) \rfloor} \tilde{m}^{\lfloor n(t \wedge N) \rfloor - j} \left(x_j - x_{j-1} + \frac{1}{n} m \alpha \right)$$

for $(x_0, x_1, \dots, x_{nN}) \in (\mathbb{R}^p)^{nN+1}$ and $t \in \mathbb{R}_+$. Then we have that $\Psi_n^N = \Psi_n^{N,2} \circ \Psi_n^{N,1}$, $n, N \in \mathbb{N}$.

Consider the set $\mathbf{C} = \{f \in \mathbf{C}(\mathbb{R}_+, \mathbb{R}^p) : \tilde{\Pi}f(0) = f(0)\}$. Let us check now that \mathbf{C} is measurable and $\mathbb{P}[\mathcal{M} \in \mathbf{C}] = 1$. The projection $\mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \ni f \mapsto \pi_0(f) := f(0) \in \mathbb{R}^p$ and the mapping $\mathbb{R}^p \ni x \mapsto (I_p - \tilde{\Pi})x \in \mathbb{R}^p$ are measurable. Since $\mathbf{C}(\mathbb{R}_+, \mathbb{R}^p)$ is a measurable set (see Ethier and Kurtz [9, Problem 3.11.25]), we have $\mathbf{C} = \mathbf{C}(\mathbb{R}_+, \mathbb{R}^p) \cap \pi_0^{-1}(\text{Null}(I_p - \tilde{\Pi})) \in \mathcal{D}_\infty(\mathbb{R}_+, \mathbb{R}^p)$. Furthermore, in Step 1, we proved that \mathcal{M} is the pathwise unique strong solution of SDE (4.4) with initial value $\mathbf{0}_p$. Hence it has continuous sample paths almost surely, and $\tilde{\Pi}\mathcal{M}_0 = \tilde{\Pi}\mathbf{0}_p = \mathbf{0}_p = \mathcal{M}_0$. Finally, the procedure to show that $\mathbf{C} \subset \mathbf{C}_\Phi, (\Phi_n)_{n \in \mathbb{N}}$ follows the same steps as on page 736 in Ispány and Pap [21].

Proof of Step 3

First, we check that $\mathcal{Z}^{(n)} = \Psi_n(\mathcal{M}^{(n)}) + \mathcal{V}^{(n)}$, $n \in \mathbb{N}$. Using Hypothesis 2 and equations (2.8), (4.1), (4.7), and (4.9), we get

$$\begin{aligned} (\Psi_n(\mathcal{M}^{(n)}))_t &= \tilde{m}^{\lfloor nt \rfloor} \frac{1}{n} \mathbf{Z}_0 + \sum_{j=1}^{\lfloor nt \rfloor} \tilde{m}^{\lfloor nt \rfloor - j} \left(\frac{1}{n} \mathbf{M}_j + \frac{1}{n} \mathbf{m}\alpha \right) \\ &= \frac{1}{n} \tilde{m}^{\lfloor nt \rfloor} \mathbf{Z}_0 + \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{m}^{\lfloor nt \rfloor - j} \left(\mathbf{Z}_j - \tilde{m} \mathbf{Z}_{j-1} - \mathbf{m}g(\mathbf{Z}_{j-1}) \right) \\ &= n^{-1} \mathbf{Z}_{\lfloor nt \rfloor} - n^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \tilde{m}^{\lfloor nt \rfloor - j} \mathbf{m}g(\mathbf{Z}_{j-1}) \\ &= \mathcal{Z}_t^{(n)} - \mathcal{V}_t^{(n)}, \quad n \in \mathbb{N}, \quad t \in \mathbb{R}_+, \end{aligned}$$

as desired.

Next, we show (4.8). Taking into account (4.5) and Jacod and Shiryaev [22, Chapter VI, Lemma 3.31], in order to prove (4.8), it is enough to see that for all $T > 0$ and $\delta > 0$, we get

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\sup_{t \in [0, T]} \|\mathcal{V}_t^{(n)}\| \geq \delta \right] = 0. \tag{4.36}$$

This can be checked as follows. For all $T > 0$ and $\delta > 0$, by Markov’s inequality and (4.29), we get

$$\begin{aligned} \mathbb{P} \left[\sup_{t \in [0, T]} \|\mathcal{V}_t^{(n)}\| \geq \delta \right] &\leq \delta^{-1} \mathbb{E} \left[n^{-1} \sum_{j=1}^{\lfloor nT \rfloor} \|\tilde{m}^{\lfloor nT \rfloor - j} \mathbf{m}g(\mathbf{Z}_{j-1})\| \right] \\ &\leq \tilde{C} \|\mathbf{m}\| \delta^{-1} n^{-1} \sum_{j=1}^{\lfloor nT \rfloor} \mathbb{E}[\|\mathbf{g}(\mathbf{Z}_{j-1})\|] \end{aligned}$$

for $n \in \mathbb{N}$, and then we proceed as in the proof of (4.27). For all $\epsilon > 0$, there exists $N(\epsilon) > 0$ such that $\|\mathbf{g}(\mathbf{z})\| < \epsilon$ for each $\mathbf{z} \in \mathbb{Z}_+^p$ with $\|\mathbf{z}\| > N(\epsilon)$. So, using Hypothesis 6 and (4.33), we get for all $T > 0$ and $\epsilon > 0$

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^{\lfloor nT \rfloor} \mathbb{E}[\|\mathbf{g}(\mathbf{Z}_{j-1})\|] &\leq \frac{1}{n} \sum_{j=1}^{\lfloor nT \rfloor} (C_g \mathbb{P}[\|\mathbf{Z}_{j-1}\| \leq N(\epsilon)] + \epsilon) \\ &\leq \epsilon T + \frac{C_g}{n} \left(\sum_{j=0}^{k_0(\epsilon, N(\epsilon)) - 1} \mathbb{P}[\|\mathbf{Z}_j\| \leq N(\epsilon)] + \sum_{j=k_0(\epsilon, N(\epsilon))}^{\lfloor nT \rfloor - 1} \epsilon \right) \\ &\leq \epsilon T + C_g k_0(\epsilon, N(\epsilon)) \frac{1}{n} + \epsilon C_g T \end{aligned} \tag{4.37}$$

for sufficiently large $n \in \mathbb{N}$. By taking the limit as $n \rightarrow \infty$, we have that for all $T > 0$ and $\epsilon > 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^{\lfloor nT \rfloor} \mathbb{E}[\|\mathbf{g}(\mathbf{Z}_{j-1})\|] \leq \epsilon T + \epsilon C_g T,$$

and then taking the limit as $\epsilon \downarrow 0$, we obtain that $n^{-1} \sum_{j=1}^{\lfloor nT \rfloor} \mathbb{E}[\|\mathbf{g}(\mathbf{Z}_{j-1})\|] \rightarrow 0$ as $n \rightarrow \infty$, yielding (4.36), as desired. \square

Appendix

The reader can consult Horn and Johnson [18] for the following known facts about primitive non-negative matrices (see Definition 8.5.0 and Theorems 8.2.8, 8.5.1 and 8.5.2 in [18]). A primitive matrix $\mathbf{A} \in \mathbb{R}_+^{p \times p}$ is an irreducible matrix with only one eigenvalue of maximum modulus (the so-called Perron–Frobenius eigenvalue), or equivalently, $\mathbf{A} \in \mathbb{R}_+^{p \times p}$ is primitive if and only if there exists $n \in \mathbb{N}$ such that $\mathbf{A}^n \in \mathbb{R}_{++}^{p \times p}$. By the Perron–Frobenius theorem, we have the following lemma.

Lemma A.1 *Let $\mathbf{A} \in \mathbb{R}_+^{p \times p}$ be a primitive matrix and let ρ be its Perron–Frobenius eigenvalue. The following statements hold:*

- (i) $\rho \in \mathbb{R}_{++}$, its algebraic and geometric multiplicity are equal to 1, and the absolute values of the other eigenvalues of \mathbf{A} are less than ρ .
- (ii) There exists a unique right eigenvector $\mathbf{u} \in \mathbb{R}_+^p$ and a unique left eigenvector $\mathbf{v} \in \mathbb{R}_+^p$ corresponding to ρ such that the sum of the coordinates of \mathbf{u} is 1 and $\mathbf{v}^\top \mathbf{u} = 1$. One calls \mathbf{u} and \mathbf{v} the right and left Perron–Frobenius eigenvector, respectively.
- (iii) If $\rho = 1$, then $\lim_{k \rightarrow \infty} \mathbf{A}^k = \Pi$ and there exist $c \in \mathbb{R}_{++}$ and $r \in (0, 1)$ such that $\|\mathbf{A}^k - \Pi\| \leq cr^k$ for each $k \in \mathbb{N}$, where $\Pi := \mathbf{u}\mathbf{v}^\top \in \mathbb{R}_{++}^{p \times p}$.

Next, we will investigate the asymptotic behaviour of the first and second moments of the norm $\|\mathbf{Z}_k\|$ for a critical CMBP $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$. We will also study the same question for the second and fourth moments of $\|\mathbf{M}_k\|$, where $(\mathbf{M}_k)_{k \in \mathbb{N}}$ is the martingale difference sequence, given in (4.1), built from $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$. But first, we need the following auxiliary result, which is easy to prove.

Lemma A.2 *Let $A, A_n, n \in \mathbb{N}$, be independent and identically distributed \mathbb{Z}_+ -valued random variables with zero mean. If B is a \mathbb{Z}_+ -valued random variable independent of A and $A_n, n \in \mathbb{N}$, then*

$$\mathbb{E} \left[\left(\sum_{i=1}^B A_i \right)^4 \right] = 3 \Sigma_A^2 (\Gamma_B + \mu_B^2) + (\zeta_A - 3 \Sigma_A^2) \mu_B,$$

where $\Sigma_A := \text{Var}[A]$, $\zeta_A := \mathbb{E}[A^4]$, $\mu_B := \mathbb{E}[B]$, and $\Gamma_B := \text{Var}[B]$.

Lemma A.3 *Let $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$ be a controlled p -type branching process given in (2.1) such that $\mathbb{E}[\|\mathbf{Z}_0\|], \mathbb{E}[\|\mathbf{X}_{0,1,i}\|]$ and $\mathbb{E}[\|\phi_0(\mathbf{z})\|]$ are finite for $i \in \{1, \dots, p\}$ and $\mathbf{z} \in \mathbb{Z}_+^p$. Assume that $\boldsymbol{\varepsilon}(\mathbf{z}) = \Lambda \mathbf{z} + \mathbf{h}(\mathbf{z})$, $\mathbf{z} \in \mathbb{Z}_+^p$, where $\Lambda \in \mathbb{R}^{p \times p}$ and $\mathbf{h} : \mathbb{Z}_+^p \rightarrow \mathbb{R}^p$ satisfies $\|\mathbf{h}(\mathbf{z})\| = \mathcal{O}(1)$ as $\|\mathbf{z}\| \rightarrow \infty$. Further, suppose that Hypothesis 5 holds.*

(i) Then we have

$$E[\|\mathbf{Z}_k\|] = O(k) \quad \text{as } k \rightarrow \infty. \tag{A.1}$$

(ii) If, in addition, $E[\|\mathbf{Z}_0\|^2]$, $E[\|\mathbf{X}_{0,1,i}\|^2]$ and $E[\|\boldsymbol{\phi}_0(\mathbf{z})\|^2]$ are finite for $i \in \{1, \dots, p\}$, $\mathbf{z} \in \mathbb{Z}_+^p$, and $\|\Gamma(\mathbf{z})\| = O(\|\mathbf{z}\|)$ as $\|\mathbf{z}\| \rightarrow \infty$, then we have (A.1) and

$$E[\|\mathbf{Z}_k\|^2] = O(k^2) \quad \text{as } k \rightarrow \infty, \tag{A.2}$$

$$E[\|\mathbf{M}_k\|^2] = O(k) \quad \text{as } k \rightarrow \infty. \tag{A.3}$$

(iii) If, in addition, Hypotheses 1, 4 and $\|\Gamma(\mathbf{z})\| = O(\|\mathbf{z}\|)$ as $\|\mathbf{z}\| \rightarrow \infty$ hold, then we have (A.1), (A.2), (A.3) and

$$E[\|\mathbf{M}_k\|^4] = O(k^2) \quad \text{as } k \rightarrow \infty. \tag{A.4}$$

Proof (i). By (2.8), for each $k \in \mathbb{N}$, we have

$$\begin{aligned} E[\mathbf{Z}_k] &= E[\mathbf{m}\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1})] = \tilde{\mathbf{m}} E[\mathbf{Z}_{k-1}] + \mathbf{m} E[\mathbf{h}(\mathbf{Z}_{k-1})] \\ &= \tilde{\mathbf{m}}^k E[\mathbf{Z}_0] + \sum_{j=0}^{k-1} \tilde{\mathbf{m}}^j \mathbf{m} E[\mathbf{h}(\mathbf{Z}_{k-1-j})]. \end{aligned} \tag{A.5}$$

Hence, using the triangle inequality, for each $k \in \mathbb{N}$, we get

$$\|E[\mathbf{Z}_k]\| \leq \|\tilde{\mathbf{m}}^k\| E[\|\mathbf{Z}_0\|] + \sum_{j=0}^{k-1} \|\tilde{\mathbf{m}}^j\| \|\mathbf{m}\| E[\|\mathbf{h}(\mathbf{Z}_{k-1-j})\|].$$

We recall the power mean inequality used in the proofs several times: for all $n \in \mathbb{N}$, $a_i \in \mathbb{R}_+$, $i = 1, \dots, n$, and $0 < k_1 \leq k_2$, we get $\left(\frac{1}{n} \sum_{i=1}^n a_i^{k_1}\right)^{\frac{1}{k_1}} \leq \left(\frac{1}{n} \sum_{i=1}^n a_i^{k_2}\right)^{\frac{1}{k_2}}$. Using $\|\mathbf{x}\| \leq \sqrt{p} \sum_{i=1}^p x_i$ for $\mathbf{x} = (x_1, \dots, x_p) \in \mathbb{R}_+^p$ and the power mean inequality, for any random vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_p)$ having non-negative coordinates, we obtain that $E[\|\boldsymbol{\xi}\|] \leq p\|E[\boldsymbol{\xi}]\|$. Consequently, we get

$$E[\|\mathbf{Z}_k\|] \leq p \|E[\mathbf{Z}_k]\| = p\tilde{C}(E[\|\mathbf{Z}_0\|] + \|\mathbf{m}\|C_h k),$$

where $\tilde{C} = \sup_{j \in \mathbb{Z}_+} \|\tilde{\mathbf{m}}^j\| < \infty$ due to Hypothesis 5 (see (4.29)) and the constant C_h is defined by

$$C_h := \sup_{\mathbf{z} \in \mathbb{Z}_+^p} \|\mathbf{h}(\mathbf{z})\| < \infty. \tag{A.6}$$

This yields (A.1).

(ii). Since the finiteness of the second moment of the norm of a random vector implies that of the first moment, using part (i) of the present lemma, we have (A.1). Taking into account that the expected value and the trace of a random square matrix commute, we have

$$E[\|\mathbf{Z}_k\|^2] = \text{tr}\left(E[\mathbf{Z}_k]E[\mathbf{Z}_k]^\top\right) + \text{tr}(\text{Var}[\mathbf{Z}_k]) \leq (E[\|\mathbf{Z}_k\|])^2 + p\|\text{Var}[\mathbf{Z}_k]\|,$$

where for the last inequality, we used that for any matrix $\mathbf{B} \in \mathbb{R}^{p \times p}$, we have $\text{tr}(\mathbf{B}) = \sum_{i=1}^p b_{i,i} \leq \sum_{i=1}^p |b_{i,i}| \leq p\|\mathbf{B}\|$. We know that (A.1) holds, therefore, in order to get (A.2), it is enough to see $\|\text{Var}[\mathbf{Z}_k]\| = O(k^2)$ as $k \rightarrow \infty$. Using the variance decomposition

formula, equations (2.8) and (2.9), and the assumption on $\boldsymbol{\varepsilon}(\mathbf{z})$, $\mathbf{z} \in \mathbb{Z}_+^p$, together with the properties of variance, for each $k \in \mathbb{N}$, we obtain

$$\begin{aligned} \text{Var}[\mathbf{Z}_k] &= \text{Var}[\mathbf{m}\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1})] + \mathbb{E}[\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1}) \odot \boldsymbol{\Sigma} + \mathbf{m}\Gamma(\mathbf{Z}_{k-1})\mathbf{m}^\top] \\ &= \text{Var}[\mathbf{m}\Lambda\mathbf{Z}_{k-1} + \mathbf{m}\mathbf{h}(\mathbf{Z}_{k-1})] + \mathbb{E}[\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1}) \odot \boldsymbol{\Sigma} + \mathbf{m}\mathbb{E}[\Gamma(\mathbf{Z}_{k-1})]\mathbf{m}^\top] \\ &= \tilde{\mathbf{m}} \text{Var}[\mathbf{Z}_{k-1}]\tilde{\mathbf{m}}^\top + \mathbf{m} \text{Var}[\mathbf{h}(\mathbf{Z}_{k-1})]\mathbf{m}^\top + \tilde{\mathbf{m}} \text{Cov}[\mathbf{Z}_{k-1}, \mathbf{h}(\mathbf{Z}_{k-1})]\mathbf{m}^\top \\ &\quad + \mathbf{m} \text{Cov}[\mathbf{h}(\mathbf{Z}_{k-1}), \mathbf{Z}_{k-1}]\tilde{\mathbf{m}}^\top + \mathbb{E}[\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1}) \odot \boldsymbol{\Sigma} + \mathbf{m}\mathbb{E}[\Gamma(\mathbf{Z}_{k-1})]\mathbf{m}^\top]. \end{aligned} \tag{A.7}$$

In what follows, we will use that for all $\mathbf{z} = (z_1, \dots, z_p) \in \mathbb{R}^p$, we have

$$\|\mathbf{z} \odot \boldsymbol{\Sigma}\| = \left\| \sum_{i=1}^p z_i \boldsymbol{\Sigma}_i \right\| \leq \|\mathbf{z}\| \sum_{i=1}^p \|\boldsymbol{\Sigma}_i\| = \|\mathbf{z}\| \|\boldsymbol{\Sigma}\|, \tag{A.8}$$

where $\|\boldsymbol{\Sigma}\| := \sum_{i=1}^p \|\boldsymbol{\Sigma}_i\|$. Proceeding recursively in (A.7), by (A.8), the triangle inequality and the symmetry of covariance, we get

$$\begin{aligned} \|\text{Var}[\mathbf{Z}_k]\| &\leq \|\tilde{\mathbf{m}}^k\|^2 \|\text{Var}[\mathbf{Z}_0]\| + \sum_{j=0}^{k-1} \|\tilde{\mathbf{m}}^j\|^2 \|\mathbf{m}\|^2 \|\text{Var}[\mathbf{h}(\mathbf{Z}_{k-1-j})]\| \\ &\quad + 2 \sum_{j=0}^{k-1} \|\tilde{\mathbf{m}}^j\|^2 \|\tilde{\mathbf{m}}\| \|\mathbf{m}\| \|\text{Cov}[\mathbf{Z}_{k-1-j}, \mathbf{h}(\mathbf{Z}_{k-1-j})]\| \\ &\quad + \sum_{j=0}^{k-1} \|\tilde{\mathbf{m}}^j\|^2 \mathbb{E}[\|\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1-j})\|] \|\boldsymbol{\Sigma}\| + \sum_{j=0}^{k-1} \|\tilde{\mathbf{m}}^j\|^2 \|\mathbf{m}\|^2 \mathbb{E}[\|\Gamma(\mathbf{Z}_{k-1-j})\|], \end{aligned} \tag{A.9}$$

and now we look for an upper bound for each term on the right hand side of (A.9).

In case of the first term, we easily have

$$\|\text{Var}[\mathbf{Z}_0]\| = \left\| \mathbb{E}[\mathbf{Z}_0(\mathbf{Z}_0)^\top] - \mathbb{E}[\mathbf{Z}_0]\mathbb{E}[(\mathbf{Z}_0)^\top] \right\| \leq \mathbb{E}[\|\mathbf{Z}_0\|^2] + (\mathbb{E}[\|\mathbf{Z}_0\|])^2 < \infty.$$

In case of the second term, since $\|\mathbf{h}(\mathbf{z})\| = O(1)$ as $\|\mathbf{z}\| \rightarrow \infty$, by (A.6), we get

$$\|\text{Var}[\mathbf{h}(\mathbf{Z}_k)]\| \leq \mathbb{E}[\|\mathbf{h}(\mathbf{Z}_k)\|^2] + (\mathbb{E}[\|\mathbf{h}(\mathbf{Z}_k)\|])^2 = O(1) \quad \text{as } k \rightarrow \infty.$$

In case of the third term, using (A.1) and (A.6), we get

$$\|\text{Cov}[\mathbf{Z}_k, \mathbf{h}(\mathbf{Z}_k)]\| \leq \mathbb{E}[\|\mathbf{Z}_k\| \|\mathbf{h}(\mathbf{Z}_k)\|] + \mathbb{E}[\|\mathbf{Z}_k\|] \mathbb{E}[\|\mathbf{h}(\mathbf{Z}_k)\|] = O(k) \quad \text{as } k \rightarrow \infty.$$

In case of the fourth term, by the assumption $\boldsymbol{\varepsilon}(\mathbf{z}) = \Lambda\mathbf{z} + \mathbf{h}(\mathbf{z})$, $\mathbf{z} \in \mathbb{Z}_+^p$, we have

$$\mathbb{E}[\|\boldsymbol{\varepsilon}(\mathbf{Z}_k)\|] \leq \|\Lambda\| \mathbb{E}[\|\mathbf{Z}_k\|] + C_h, \quad k \in \mathbb{N},$$

where the constant C_h is defined in (A.6). Using (A.1), this implies that

$$\mathbb{E}[\|\boldsymbol{\varepsilon}(\mathbf{Z}_k)\|] = O(k) \quad \text{as } k \rightarrow \infty. \tag{A.10}$$

In case of the fifth term, by the assumption $\|\Gamma(\mathbf{z})\| = O(\|\mathbf{z}\|)$ as $\|\mathbf{z}\| \rightarrow \infty$ and the equation (A.1), we get

$$\mathbb{E}[\|\Gamma(\mathbf{Z}_k)\|] \leq \|\Gamma(\mathbf{0}_p)\| + C_\Gamma \mathbb{E}[\|\mathbf{Z}_k\|] = O(k) \quad \text{as } k \rightarrow \infty, \tag{A.11}$$

where

$$C_\Gamma = \sup_{z \in \mathbb{Z}_+^p \setminus \{0_p\}} \|z\|^{-1} \|\Gamma(z)\| < \infty.$$

Taking into account the previous estimations and that $\sup_{j \in \mathbb{Z}_+} \|\tilde{m}^j\| < \infty$ due to Hypothesis 5 (see (4.29)), the inequality (A.9) implies that

$$\|\text{Var}[\mathbf{Z}_k]\| = O(1) + O(k) + \sum_{j=0}^{k-1} O(k-1-j) = O(k^2) \quad \text{as } k \rightarrow \infty,$$

which concludes the proof of (A.2).

Next, we verify (A.3). Then

$$E[\|\mathbf{M}_k\|^2] = E\left[\text{tr}\left(\mathbf{M}_k(\mathbf{M}_k)^\top\right)\right] = \text{tr}\left(E\left[\mathbf{M}_k(\mathbf{M}_k)^\top\right]\right), \quad k \in \mathbb{N}, \tag{A.12}$$

where, by the tower rule, we have $E[\mathbf{M}_k(\mathbf{M}_k)^\top] = E\left[\text{Var}\left[\mathbf{Z}_k \mid \mathcal{F}_{k-1}\right]\right]$. Consequently, using (2.9), (A.8) and the inequality $\text{tr}(\mathbf{B}) \leq p\|\mathbf{B}\|$ for any matrix $\mathbf{B} \in \mathbb{R}^{p \times p}$ (justified earlier), we get that

$$\begin{aligned} \text{tr}\left(E\left[\mathbf{M}_k(\mathbf{M}_k)^\top\right]\right) &= \text{tr}\left(E[\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1})] \odot \boldsymbol{\Sigma}\right) + \text{tr}\left(\mathbf{m}E[\Gamma(\mathbf{Z}_{k-1})]\mathbf{m}^\top\right) \\ &\leq p\|E[\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1})] \odot \boldsymbol{\Sigma}\| + p\|\mathbf{m}E[\Gamma(\mathbf{Z}_{k-1})]\mathbf{m}^\top\| \\ &\leq pE[\|\boldsymbol{\varepsilon}(\mathbf{Z}_{k-1})\|]\|\boldsymbol{\Sigma}\| + pE[\|\Gamma(\mathbf{Z}_{k-1})\|]\|\mathbf{m}\|^2, \quad k \in \mathbb{N}. \end{aligned} \tag{A.13}$$

Therefore, (A.3) follows from (A.10), (A.11), (A.12) and (A.13).

(iii). Consider the following reformulation of the martingale difference sequence $(\mathbf{M}_k)_{k \in \mathbb{N}}$ defined in (4.1):

$$\mathbf{M}_k = \sum_{i=1}^p \left(\sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} \mathbf{X}_{k-1,j,i} - \varepsilon_i(\mathbf{Z}_{k-1})\mathbf{m}_i \right), \quad k \in \mathbb{N},$$

where we used (2.8).

Applying the power mean inequality twice, and adding and subtracting the random variable $\sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} \mathbf{m}_i = \phi_{k-1,i}(\mathbf{Z}_{k-1})\mathbf{m}_i$, we get

$$\begin{aligned} \|\mathbf{M}_k\|^4 &\leq \left(\sum_{i=1}^p \left\| \sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} \mathbf{X}_{k-1,j,i} - \varepsilon_i(\mathbf{Z}_{k-1})\mathbf{m}_i \right\| \right)^4 \\ &\leq p^3 \sum_{i=1}^p \left\| \sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} \mathbf{X}_{k-1,j,i} - \varepsilon_i(\mathbf{Z}_{k-1})\mathbf{m}_i \right\|^4 \\ &\leq 8p^3 \sum_{i=1}^p \left(\left\| \sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} (\mathbf{X}_{k-1,j,i} - \mathbf{m}_i) \right\|^4 + \|(\phi_{k-1,i}(\mathbf{Z}_{k-1}) - \varepsilon_i(\mathbf{Z}_{k-1}))\mathbf{m}_i\|^4 \right) \end{aligned}$$

for $k \in \mathbb{N}$. Using again the power mean inequality and that $\|\mathbf{z}\|^4 = (z_1^2 + \dots + z_p^2)^2$, $\mathbf{z} = (z_1, \dots, z_p) \in \mathbb{Z}_+^p$, we get

$$\begin{aligned} \mathbb{E}[\|\mathbf{M}_k\|^4] &\leq 8p^4 \sum_{i=1}^p \sum_{l=1}^p \mathbb{E} \left[\left(\sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} (\mathbf{X}_{k-1,j,i,l} - m_{i,l}) \right)^4 \right] \\ &\quad + 8p^4 \sum_{i=1}^p \sum_{l=1}^p \mathbb{E} \left[((\phi_{k-1,i}(\mathbf{Z}_{k-1}) - \varepsilon_i(\mathbf{Z}_{k-1}))m_{i,l})^4 \right]. \end{aligned}$$

We will compute the previous upper bounds by first determining the corresponding conditional expectations with respect to \mathcal{F}_{k-1} . Using the notations (2.7) and the Markov property of $(\mathbf{Z}_k)_{k \in \mathbb{Z}_+}$, for each $k \in \mathbb{N}$, $i, l \in \{1, \dots, p\}$, we obtain

$$\mathbb{E} \left[(\phi_{k-1,i}(\mathbf{Z}_{k-1}) - \varepsilon_i(\mathbf{Z}_{k-1}))^4 m_{i,l}^4 \mid \mathcal{F}_{k-1} \right] = m_{i,l}^4 \kappa_i(\mathbf{Z}_{k-1}),$$

and, by (2.6), (2.7) and Lemma A.2 together with the independence of $\phi_{k-1}(\mathbf{z})$, $\mathbf{z} \in \mathbb{Z}_+^p$, $\mathbf{X}_{k-1,j,i}$, $j \in \mathbb{N}$, $i \in \{1, \dots, p\}$ and \mathbf{Z}_{k-1} , we get that

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{j=1}^{\phi_{k-1,i}(\mathbf{Z}_{k-1})} (\mathbf{X}_{k-1,j,i,l} - m_{i,l}) \right)^4 \mid \mathcal{F}_{k-1} \right] &= 3\Sigma_{i,l,l}^2 (\Gamma_{i,i}(\mathbf{Z}_{k-1}) + \varepsilon_i(\mathbf{Z}_{k-1})^2) \\ &\quad + (\zeta_{i,l} - 3\Sigma_{i,l,l}^2)\varepsilon_i(\mathbf{Z}_{k-1}), \end{aligned}$$

where $\Sigma_{i,l,l}$ is the l -th diagonal element of the variance matrix Σ_i (see (2.6)).

Consequently, we get

$$\begin{aligned} \mathbb{E}[\|\mathbf{M}_k\|^4] &\leq 8p^4 \sum_{i=1}^p \sum_{l=1}^p \left(m_{i,l}^4 \mathbb{E}[\kappa_i(\mathbf{Z}_{k-1})] + 3\Sigma_{i,l,l}^2 \mathbb{E}[\Gamma_{i,i}(\mathbf{Z}_{k-1})] \right. \\ &\quad \left. + 3\Sigma_{i,l,l}^2 \mathbb{E}[\varepsilon_i(\mathbf{Z}_{k-1})^2] + (\zeta_{i,l} - 3\Sigma_{i,l,l}^2) \mathbb{E}[\varepsilon_i(\mathbf{Z}_{k-1})] \right). \end{aligned}$$

Thus, to get (A.4) it is enough to check that for each $i \in \{1, \dots, p\}$, $\mathbb{E}[\varepsilon_i(\mathbf{Z}_k)]$, $\mathbb{E}[\Gamma_{i,i}(\mathbf{Z}_k)]$, $\mathbb{E}[\varepsilon_i(\mathbf{Z}_k)^2]$, $\mathbb{E}[\kappa_i(\mathbf{Z}_k)]$ are $\mathcal{O}(k^2)$ as $k \rightarrow \infty$. Using (A.10) and (A.11), we get

$$\mathbb{E}[\varepsilon_i(\mathbf{Z}_k)] \leq \mathbb{E}[\|\boldsymbol{\varepsilon}(\mathbf{Z}_k)\|] = \mathcal{O}(k), \quad \mathbb{E}[\Gamma_{i,i}(\mathbf{Z}_k)] \leq \mathbb{E}[\|\Gamma(\mathbf{Z}_k)\|] = \mathcal{O}(k)$$

as $k \rightarrow \infty$ for $i \in \{1, \dots, p\}$. Moreover, by (A.1) and (A.2), we obtain

$$\mathbb{E}[\varepsilon_i(\mathbf{Z}_k)^2] \leq \mathbb{E}[\|\boldsymbol{\varepsilon}(\mathbf{Z}_k)\|^2] = \mathbb{E}[\|\Lambda \mathbf{Z}_k + \mathbf{h}(\mathbf{Z}_k)\|^2] \leq 2(\|\Lambda\|^2 \mathbb{E}[\|\mathbf{Z}_k\|^2] + C_{\mathbf{h}}^2) = \mathcal{O}(k^2)$$

as $k \rightarrow \infty$ for $i \in \{1, \dots, p\}$, where C_h is defined in (A.6). Finally, by Hypothesis 4, we get

$$C_k := \sup_{\substack{z \in \mathbb{Z}_+^p \setminus \{\mathbf{0}_p\} \\ i=1, \dots, p}} \|z\|^{-2} \|\kappa_i(z)\| < \infty,$$

and hence, by (A.2), we have $E[\kappa_i(\mathbf{Z}_k)] \leq \kappa_i(\mathbf{0}_p) + C_k E[\|\mathbf{Z}_k\|^2] = O(k^2)$ as $k \rightarrow \infty$ for $i \in \{1, \dots, p\}$. □

Next, we recall a result on weak convergence of random step processes toward a diffusion process due to Ispány and Pap [20, Corollary 2.2].

Theorem A.4 *Let $\mathbf{b} : \mathbb{R}_+ \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ and $\mathbf{C} : \mathbb{R}_+ \times \mathbb{R}^p \rightarrow \mathbb{R}^{p \times r}$ be continuous functions. Assume that uniqueness in the sense of probability law holds for the SDE*

$$d\mathbf{U}_t = \mathbf{b}(t, \mathbf{U}_t) dt + \mathbf{C}(t, \mathbf{U}_t) d\mathbf{W}_t, \quad t \in \mathbb{R}_+, \tag{A.14}$$

with initial value $\mathbf{U}_0 = \mathbf{u}_0$ for all $\mathbf{u}_0 \in \mathbb{R}^p$, where $(\mathbf{W}_t)_{t \in \mathbb{R}_+}$ is an r -dimensional standard Wiener process. Let η be a probability measure on $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$, and let $(\mathbf{U}_t)_{t \in \mathbb{R}_+}$ be a solution of (A.14) with initial distribution η . For each $n \in \mathbb{N}$, let $(\mathbf{U}_k^{(n)})_{k \in \mathbb{Z}_+}$ be a sequence of \mathbb{R}^p -valued random vectors adapted to a filtration $(\mathcal{F}_k^{(n)})_{k \in \mathbb{Z}_+}$ (i.e. $\mathbf{U}_k^{(n)}$ is $\mathcal{F}_k^{(n)}$ -measurable) such that $E[\|\mathbf{U}_k^{(n)}\|^2] < \infty$ for each $n, k \in \mathbb{N}$. Let

$$\mathbf{U}_t^{(n)} := \sum_{k=0}^{\lfloor nt \rfloor} \mathbf{U}_k^{(n)}, \quad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Suppose $\mathbf{U}_0^{(n)} \xrightarrow{\mathcal{L}} \eta$ as $n \rightarrow \infty$, and that for all $T > 0$,

- (i) $\sup_{t \in [0, T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} E \left[\mathbf{U}_k^{(n)} \mid \mathcal{F}_{k-1}^{(n)} \right] - \int_0^t \mathbf{b}(s, \mathbf{U}_s^{(n)}) ds \right\| \xrightarrow{\mathbb{P}} 0$ as $n \rightarrow \infty$,
- (ii) $\sup_{t \in [0, T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \text{Var} \left[\mathbf{U}_k^{(n)} \mid \mathcal{F}_{k-1}^{(n)} \right] - \int_0^t \mathbf{C}(s, \mathbf{U}_s^{(n)}) \mathbf{C}(s, \mathbf{U}_s^{(n)})^\top ds \right\| \xrightarrow{\mathbb{P}} 0$ as $n \rightarrow \infty$,
- (iii) $\sum_{k=1}^{\lfloor nT \rfloor} E \left[\|\mathbf{U}_k^{(n)}\|^2 \mathbf{1}_{\{\|\mathbf{U}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1}^{(n)} \right] \xrightarrow{\mathbb{P}} 0$ as $n \rightarrow \infty$ for all $\theta > 0$.

Then $\mathbf{U}^{(n)} \xrightarrow{\mathcal{L}} \mathbf{U}$ as $n \rightarrow \infty$.

For measurable mappings $\Phi, \Phi_n : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $n \in \mathbb{N}$, let $\mathbf{C}_{\Phi, (\Phi_n)_{n \in \mathbb{N}}}$ be the set of all functions $\mathbf{f} \in \mathbf{C}(\mathbb{R}_+, \mathbb{R}^p)$ for which $\Phi_n(\mathbf{f}_n) \rightarrow \Phi(\mathbf{f})$ whenever $\mathbf{f}_n \xrightarrow{\text{lu}} \mathbf{f}$ with $\mathbf{f}_n \in \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $n \in \mathbb{N}$. The notation $\xrightarrow{\text{lu}}$ refers to local uniform convergence, i.e., $\sup_{t \in [0, T]} \|\mathbf{f}_n(t) - \mathbf{f}(t)\| \rightarrow 0$ as $n \rightarrow \infty$ for all $T \in \mathbb{R}_{++}$. The following result is a kind of continuous mapping theorem, which can be considered as a consequence of Theorem 3.27 in Kallenberg [24], and its proof also appears in Ispány and Pap [20, Lemma 3.1].

Theorem A.5 *Let $(\mathbf{U}_t)_{t \in \mathbb{R}_+}$ and $(\mathbf{U}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, be \mathbb{R}^p -valued stochastic processes with càdlàg paths such that $\mathbf{U}^{(n)} \xrightarrow{\mathcal{L}} \mathbf{U}$ as $n \rightarrow \infty$. Let $\Phi : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$ and $\Phi_n : \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p) \rightarrow \mathbf{D}(\mathbb{R}_+, \mathbb{R}^p)$, $n \in \mathbb{N}$, be measurable mappings such that there exists a measurable set $C \subset \mathbf{C}_{\Phi, (\Phi_n)_{n \in \mathbb{N}}}$ with $\mathbb{P}[\mathbf{U} \in C] = 1$. Then $\Phi_n(\mathbf{U}^{(n)}) \xrightarrow{\mathcal{L}} \Phi(\mathbf{U})$ as $n \rightarrow \infty$.*

Acknowledgements We would like to thank the referee for her/his comments that helped us improve the paper.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. Mátyás Barczy was supported by the project TKP2021-NVA-09. Project no. TKP2021-NVA-09 has been implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme. Miguel González, Pedro Martín-Chávez and Inés del Puerto are supported by grant PID2019-108211GB-I00 funded by MICIU/AEI/10.13039/501100011033. Pedro Martín-Chávez is also grateful to the Spanish Ministry of Universities for support from a predoctoral fellowship Grant no. FPU20/06588.

Data availability The authors declare that they did not use any data set in their paper.

Declarations

Conflicts of interest The authors declare that they have no Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alsmeyer, G., Rösler, U.: The bisexual Galton-Watson process with promiscuous mating: extinction probabilities in the supercritical case. *Ann. Appl. Probab.* **6**(3), 922–939 (1996)
2. Alsmeyer, G., Rösler, U.: Asexual versus promiscuous bisexual Galton-Watson processes: The extinction probability ratio. *Ann. Appl. Probab.* **12**(1), 125–142 (2002)
3. Asmussen, S.: On some two-sex population models. *Ann. Probab.* **8**(4), 727–744 (1980)
4. Athreya, K.B., Ney, P.E.: *Branching Processes*. 1st ed. In: *Grundlehren der mathematischen Wissenschaften*, 1st edn. Springer, Berlin, Heidelberg 1972
5. Bansaye, V., Caballero, M., Méléard, S., San Martín, J.: Scaling limits of bisexual Galton-Watson processes. *Stochastics* **95**(5), 749–784 (2023)
6. Barczy, M., Ispány, M., Pap, G.: Asymptotic behavior of unstable INAR(p) processes. *Stoch. Process. Appl.* **121**(3), 583–608 (2011)
7. Barczy, M., Pap, G.: Asymptotic behavior of critical, irreducible multi-type continuous state and continuous time branching processes with immigration. *Stoch. Dyn.* **16**, 04 (2016). Article 1650008
8. Daley, D.J.: Extinction conditions for certain bisexual Galton-Watson branching processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **9**, 315–322 (1968)
9. Ethier, S.N., Kurtz, T.G.: *Markov Processes: Characterization and Convergence*. In: *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York (1986)
10. Feller, W.: *Diffusion processes in genetics*. In: Neyman, K. (ed.) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, vol. 2 of *Berkeley Symposium on Mathematical Statistics and Probability*, pp. 227–246. University of California Press, Berlin, Heidelberg (1951)
11. González, M., del Puerto, I.: Weighted conditional least squares estimation in controlled multitype branching processes. In: González Velasco, M., del Puerto, I., Martínez, R., Molina, M., Mota, M., Ramos, A. (eds.) *Workshop on Branching Processes and Their Applications*, pp. 147–155. Springer, Berlin Heidelberg (2010)
12. González, M., del Puerto, I., Yanev, G.: *Controlled branching processes*. ISTE, Wiley, London (2018)
13. González, M., Martín-Chávez, P., del Puerto, I.: Diffusion approximation of controlled branching processes using limit theorems for random step processes. *Stoch. Model.* **39**(1), 232–248 (2023)

14. González, M., Martínez, R., Mota, M.: On the unlimited growth of a class of homogeneous multitype Markov chains. *Bernoulli* **11**(3), 559–570 (2005)
15. González, M., Minuesa, C., del Puerto, I., Vidyashankar, A.N.: Robust Estimation in Controlled Branching Processes: Bayesian Estimators via Disparities. *Bayesian Anal.* **16**(3), 1009–1037 (2021)
16. González, M., Martínez, R., Mota, M.: Rates of growth in a class of homogeneous multidimensional Markov chains. *J. Appl. Probab.* **43**(1), 159–174 (2006)
17. González, M., Molina, M., del Puerto, I.: Asymptotic behaviour of critical controlled branching processes with random control functions. *J. Appl. Probab.* **42**(2), 463–477 (2005)
18. Horn, R.A., Johnson, C.R.: *Matrix Analysis*, 2nd edn. Cambridge University Press (1985)
19. Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, 2nd edn. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo (1989)
20. Ispány, M., Pap, G.: A note on weak convergence of random step processes. *Acta Math. Hung.* **126**, 381–395 (2010)
21. Ispány, M., Pap, G.: Asymptotic behavior of critical primitive multi-type branching processes with immigration. *Stoch. Anal. Appl.* **32**(5), 727–741 (2014)
22. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, Heidelberg (2003)
23. Jagers, P.: The proportions of individuals of different kinds in two-type populations. a branching process problem arising in biology. *J. Appl. Probab.* **2**, 249–260 (1969)
24. Kallenberg, O.: *Foundations of modern probability*, 1st edn. In: *Probability and Its Applications*. Springer, New York, NY (1997)
25. Kaplan, N.: The multitype Galton-Watson process with immigration. *Ann. Probab.* **1**(6), 947–953 (1973)
26. Kimmel, M., Axelrod, D.: *Branching processes in biology*, 2nd edn. Springer, New York, NY (2015)
27. Mwasunda, J.A., Irunde, J.I., Kajunguri, D., Kuznetsov, D.: Outbreak or extinction of bovine cysticercosis and human taeniasis: A stochastic modelling approach. *Appl. Math. Model.* **106**, 73–85 (2022)
28. Ráth, B.: Time evolution of dense multigraph limits under edge-conservative preferential attachment dynamics. *Random Struct. Algorithms* **41**(3), 365–390 (2012)
29. Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion*, 3rd edn. Grundlehren der mathematischen Wissenschaften. Springer, Berlin, Heidelberg (1998)
30. Sagitov, S., StÅhlberg, A.: Counting unique molecular identifiers in sequencing using a multi-type branching process with immigration. *J. Theor. Biol.* **558**, 111365 (2023)
31. Sevast'yanov, B.A., Zubkov, A.M.: Controlled branching processes. *Theory Prob. Appl.* **19**(1), 14–24 (1974)
32. Sriram, T., Bhattacharya, A., González, M., Martínez, R., del Puerto, I.: Estimation of the offspring mean in a controlled branching process with a random control function. *Stoch. Process. Appl.* **117**(7), 928–946 (2007)
33. Wei, C.Z., Winnicki, J.: Some asymptotic results for the branching process with immigration. *Stoch. Process. Appl.* **31**(2), 261–282 (1989)
34. Yakovlev, A.Y., Yanev, N.M.: Relative frequencies in multitype branching processes. *Ann. Appl. Probab.* **19**(1), 1–14 (2009)
35. Yakovlev, A.Y., Yanev, N.M.: Limiting distributions for multitype branching processes. *Stoch. Anal. Appl.* **28**(6), 1040–1060 (2010)
36. Yamada, T., Watanabe, S.: On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11**(1), 155–167 (1971)
37. Yanev, N.M.: Stochastic models of cell proliferation kinetics based on branching processes. In: Almudevar, A., Oakes, D., Hall, J. (eds.) *Stat. Model. Biol. Syst. Memory Andrei Yakovlev*, pp. 3–20. Springer International Publishing (2020)