Stochastic Processes and their Applications 41 (1992) 297-324
North-Holland

A non-Markovian model for cell population growth: Speed of convergence and central limit theorem

Mathisca C.M. de Gunst*
Department of Mathematics, Free University of Amsterdam, Netherlands

Willem R. van Zwet**
Department of Mathematics, University of Leiden, Netherlands, and Department of Statistics, University of North Carolina, Chapel Hill, NC, USA

Received 13 October 1989
Revised 28 February 1991

In De Gunst (1989) a stochastic model was developed for the growth of a batch culture of plant cells. In this paper the mathematical properties of the model are considered. We investigate the asymptotic behaviour of the population growth as predicted by the model when the initial cell number of population members tends to infinity. In particular it is shown that the total cell number, which is a non-Markovian counting process, converges almost surely, uniformly on the real line to a non-random function and the rate of convergence is established. Moreover, a central limit theorem is proved. Computer simulations illustrate the behaviour of the process. The model is graphically compared with experimental data.

AMS 1980 Subject Classifications: Primary 60G55, 60F05, 60F15; Secondary 62P10.

stochastic model * population growth * non-Markovian counting process * almost sure convergence * rate of convergence * central limit theorem

1. Introduction

Studies of the growth of plant cell cultures in liquid medium raised several questions which could not be answered by means of the available techniques. This led to the formulation of a stochastic model for plant cell population growth (De Gunst, 1989). The model arises naturally from hypotheses concerning the underlying biological mechanisms, i.e. it starts with a mathematical description of the individual cell

---

Correspondence to: Dr. M.C.M. de Gunst, Department of Mathematics, Free University of Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, Netherlands.

* Supported in part by a NATO Science Fellowship from the Netherlands Organization for Scientific Research (NWO).

** Supported in part by the Office of Naval Research Contract N0014-80-C-1063.

behaviour based on present biological knowledge. The model has turned out to give a good description of actual plant cell population growth and to predict some interesting features concerning the growth which were unknown before.

In this paper we consider the mathematical properties of the model. In particular we investigate the asymptotic behaviour of the population growth as predicted by the model when the initial number of population members tends to infinity. A detailed description of the biological background, experimental procedures, and statistical analysis of the data can be found in De Gunst (1989) and De Gunst et al. (1990).

Although this type of model is not necessarily restricted to a description of plant cell population growth, but may be used to describe population behaviour when the growth process is bounded and shows a non-Markovian behaviour in general, to fix our thoughts we prefer to formulate the model in the context of this example. In Section 2 we give a brief description of the biological phenomenon. The model is formulated in Section 3. The next two sections deal with the asymptotic population behaviour: in Section 4 it is shown that the total cell number converges almost surely uniformly on the whole real (half-)line to a non-random function and the rate of convergence is derived; in Section 5 a central limit theorem is proved. In a companion paper (De Gunst and Van Zwet, to appear) we investigate the tail behaviour and duration of the growth process.

To see what the predicted cell number looks like for the proposed stochastic process and its limit function, we have simulated the stochastic growth curve and computed the limit function numerically for several sets of parameter values. The resulting curves are shown in Section 6, as well as the results of one of the experiments which were performed in order to compare the model with real data.

2. Biological background

The phenomenon of interest is the growth of a so-called batch culture of plant cells. A batch culture is a culture of isolated plant cells or very small cell aggregates which remain dispersed as they grow in a liquid medium. Usually the cells grow in an Erlenmeyer flask which is continuously shaken or in a fermentor which is continuously stirred. In contrast to a continuous culture, a batch culture neither has an inflow of fresh medium nor an outflow of culture (Street, 1973).

At some time, which will be called time $t=0$, the cells, say n in number, are transferred to a fresh medium of known composition. After a certain time period, the so-called lag phase, the first cell divisions are observed. The number of cells is seen to increase until the population seems to have stopped growing, probably because some components in the medium have run low. When the cells are transferred to a fresh medium again, this growth pattern repeats itself.

The cultures considered seem to consist of two types of cells: type-A cells, which are actively cycling and finally divide, and type-B cells, which are 'resting' or differentiating and don't divide. At time $t = 0$, i.e. just after the transfer to a fresh medium, the population is assumed to consist of type-A cells only. Since cell death other than division does not seem to play a substantial role during the growth process of the population, it is neglected in the sequel.

The time between birth and division of type-A cells living under equal external circumstances is observed to be of variable duration (see for example King and Street, 1973). It is believed that a random time after its birth a type-A cell receives an endogenous stimulus which starts some biochemical processes and that an approximately constant time (c) is needed for these processes to produce the cell's division. Thus the A-cells can be divided into two classes, to wit A^0-cells which are waiting for a stimulus, and A^1-cells which have already received a stimulus, but have not yet divided. In general the total number of cells in a culture can be experimentally established, but there is no way yet to distinguish A-cells from B-cells unambiguously, so that no quantitative data are available for the different cell types.

At least two components in the medium are observed to play a role in the regulation of cell division: sugar, which is the main substrate for building up new cell material, and hormone. The length of the time period between birth and receipt of stimulus is thought to depend on the substrate concentration in the medium: the higher the concentration, the shorter the mean length. It is assumed that a fixed amount (y_s^{-1}) of substrate is consumed when a cell receives a stimulus, and as a result this mean length becomes longer when time progresses. When all substrate is used up, no further stimuli will be received, so that there will be no more divisions.

The amount of hormone in the medium seems to influence (future) cell division. For that reason its concentration is assumed to determine whether a cell just being born will be a cycling A-cell or will differentiate. The higher the hormone concentration, the more A-cells will be born. Also the amount of hormone is supposed to decrease with fixed amounts (y_h^{-1}), but now at each time a division takes place. When the hormone has run out, only B-cells are born.

These are the main biological features on which the model as described in the next section is based.

3. The model

For $n = 1, 2, \ldots$ we consider a model starting with $N_n(0) = N_{An}(0) = n$ cells in the A-state. The total number of cells at time t and the number of type-A cells at time t are denoted by $N_n(t)$ and $N_{An}(t)$, respectively. In practice the cell numbers are usually very large (10^5–10^6 ml^{-1}). For this reason we investigate the population behaviour as the number of individuals tends to infinity. The obvious quantity to which the number of individuals is related is the initial cell number n, which is also usually very large in practice. Hence we shall commit ourselves to asymptotics in n. This is why all random variables in the model are given the index n. Note that in this context a quantity being fixed or constant means that it does not depend on

n; it may of course depend on the type of plant cells used for the experiments and/or be unknown in practice.

The growth of the whole cell population in the model is determined by the behaviour of the individual cells, which is assumed to be ruled by the following regime.

Fix c, independent of n, and define the time of the ith division by

$$\tau_{in} = \inf\{t\colon N_n(t) = n + i\}, \quad i = 1, 2, \ldots. \tag{3.1}$$

A cell dividing at time τ_{in} has started its mitotic cycle at time $\tau_{in} - c$, because it received a stimulus at this time. A type-A cell receives this stimulus at the time of the first event in a counting process starting at the birth of the cell and having a variable, random rate $Q_n(t)$. For different cells the processes are coupled through this rate, but otherwise independent.

In providing a stimulus a fixed quantity y_s^{-1} of substrate is used up. The amount of substrate $S_n(t)$ at time t is given by

$$S_n(t) = \begin{cases} S_n(0) = nb_s y_s^{-1}, & 0 \leqslant t < \tau_{1n} - c, \\ S_n(0) - iy_s^{-1} = (nb_s - i)y_s^{-1}, & \tau_{in} - c \leqslant t < \tau_{i+1,n} - c, \\ & i = 1, 2, \ldots, n_{sn} - 1, \\ 0, & t \geqslant \tau_{n_{sn}n} - c, \end{cases} \tag{3.2}$$

where b_s is fixed and $n_{sn} = nb_s$. Note that for convenience we have assumed that $nb_s \in \mathbb{N}$. Although this is not necessarily true in practice, all proofs are easily extended to include the case $nb_s \notin \mathbb{N}$ too (see De Gunst, 1989).

It is assumed that the rate $Q_n(t)$ depends on $S_n(t)$ in the following way.

$$Q_n(t) = \frac{S_n(t)}{d(S_n(t) + nk_s)}, \quad t \geqslant 0,$$

$$= \begin{cases} \dfrac{nb_s}{d(nb_s + nk_s y_s)} \equiv Q_{0n} \equiv Q_0, & 0 \leqslant t < \tau_{1n} - c, \\ \dfrac{nb_s - i}{d(nb_s + nk_s y_s - i)} \equiv Q_{in}, & \tau_{in} - c \leqslant t < \tau_{i+1,n} - c, \\ & i = 1, 2, \ldots, n_{sn} - 1, \\ 0 \equiv Q_{in}, & t \geqslant \tau_{n_{sn}n} - c, \\ & i = n_{sn}, n_{sn} + 1, \ldots, \end{cases} \tag{3.3}$$

where d and k_s are constants. This means that the total number of stimuli is never larger than n_{sn}, since $Q_n(t) = 0$ for $t \geqslant \tau_{n_{sn}n} - c$.

In the model the amount of hormone is given by

$$H_n(\tau_{in}) = \begin{cases} H_n(0) = nb_h y_h^{-1}, & i = 1, \\ H_n(0) - (i-1)y_h^{-1} = (nb_h - i + 1)y_h^{-1}, & i = 2, 3, \ldots, n_{hn}, \\ 0, & i = n_{hn} + 1, n_{hn} + 2, \ldots, \end{cases} \tag{3.4}$$

where y_{h} and b_{h} are independent of n, $n_{\mathrm{h}n} = nb_{\mathrm{h}}$ and again we have assumed without loss of generality that $nb_{\mathrm{h}} \in \mathbb{N}$.

The two cells originating at time τ_{in} independently become type-A cells with probability P_{in} and type-B cells with probability $1 - P_{in}$, where P_{in} depends on the amount of hormone present:

$$P_{in} = \frac{H_n(\tau_{in})}{H_n(\tau_{in}) + nk_{\mathrm{h}}}, \quad i = 1, 2, \ldots,$$

$$= \begin{cases} \dfrac{nb_{\mathrm{h}} - i + 1}{nb_{\mathrm{h}} + nk_{\mathrm{h}}y_{\mathrm{h}} - i + 1}, & i = 1, 2, \ldots, n_{\mathrm{h}n}, \\ 0, & i = n_{\mathrm{h}n} + 1, n_{\mathrm{h}n} + 2, \ldots, \end{cases} \tag{3.5}$$

where k_{h} is fixed. Note that the processes may continue after $n_{\mathrm{h}n}$ divisions, even though $P_{n_{\mathrm{h}n}+k,n} = 0$, $k = 1, 2, \ldots$. Then only type-B cells are produced. Finally, we should point out that the way we have assumed the substrate and the hormone to act (cf. (3.3) and (3.5)) is derived from the well-known Monod kinetics (see for example Roels, 1983).

Having formulated the model on the individual level, we now consider what this means for the total population behaviour.

Let $t \geq c$. Before or at time $t - c$, $N_n(t - c) - n$ type-A cells have divided. Therefore the number of type-A cells at time $t - c$ is given by

$$N_{\mathrm{A}n}(t - c) = 2n - N_n(t - c) + \sum_{i=1}^{N_n(t-c)-n} Z_{in}, \tag{3.6}$$

where Z_{in} is the number of newly born A-cells at the ith division. Thus $Z_{1n}, Z_{2n}, \ldots$ are independent random variables and Z_{in} has a binomial distribution with parameters 2 and P_{in}, $i = 1, 2, \ldots$. Since $P_{in} = 0$ for $i > n_{\mathrm{h}n}$, $Z_{in} = 0$ a.s. for $i > n_{\mathrm{h}n}$. We shall find it convenient to write

$$\hat{Z}_n = (Z_{1n}, Z_{2n}, \ldots), \qquad Z_n(t) = \sum_{i=1}^{N_n(t)-n} Z_{in}.$$

Let the number of type-A cells waiting for a stimulus at time t, i.e. the type-A cells which have not received a stimulus before t, be denoted by $N^0_{\mathrm{A}n}(t)$. Then, since at time $t - c$ there are $N_n(t) - N_n(t - c)$ A-cells which have received a stimulus and have not yet divided,

$$N^0_{\mathrm{A}n}(t - c) = 2n - N_n(t) + Z_n(t - c). \tag{3.7}$$

Since at time $t - c$ all of these $N^0_{\mathrm{A}n}(t - c)$ cells are waiting for stimuli arriving independently with rate $Q_n(t - c)$, and a cell divides time c after receiving a stimulus, the process $N_n(t)$ grows at a variable random rate $N^0_{\mathrm{A}n}(t - c)Q_n(t - c)$. The two factors entering into this rate are given by (3.3) and (3.7). Note that the observed process $N_n(t)$ together with the observed random vector $\hat{Z}_n$ determine all other random quantities discussed so far.

The process $N_n(t)$ stops at the first time when either $N_{An}(t)=0$ or $N_n(t)=n+n_{sn}$. Note that $N^0_{An}(t-c)=0$ is not sufficient for N_n to stop growing at time t, since new A-cells may be born between time $t-c$ and t. Thus from (3.6) we see that the process stops at the random time T_n defined by

$$T_n=\inf\left\{t:\ \sum_{i=1}^{N_n(t)-n}(Z_{in}-1)=-n\right\}\wedge\inf\{t:\ N_n(t)=n+n_{sn}\},\tag{3.8}$$

where $a\wedge b$ means the smaller of a and b. The random level which $N_n(t)$ has reached by then is

$$N_n(T_n)=n+\inf\left\{m:\ \sum_{i=1}^{m}(Z_{in}-1)=-n\right\}\wedge n_{sn},\tag{3.9}$$

and we see that given $\hat{Z}_n$, this final level is a non-random quantity.

In what follows we assume that $\{N_n(t)\}_{t\geqslant 0}$ and $\hat{Z}_n$ for $n=1,2,\ldots$ are defined on a single probability space $(\Omega,\mathscr{G},\mathscr{P})$ such that $Z_{1n},Z_{2n},\ldots$ are independent and Z_{in} has a binomial distribution with parameters 2 and P_{in} given by (3.5), and conditionally on $\hat{Z}_n$, $\{N_n(t)-n\}_{t\geqslant 0}$ is distributed under $\mathscr{P}$ as a stopped non-Markovian counting process with intensity process $\Lambda_{zn}(t)$ given by the left-continuous version of

$$\begin{aligned}\Lambda_{zn}(t)&=\lim_{h\downarrow 0}h^{-1}\mathscr{P}(N_n(t+h)-N_n(t)\geqslant 1\,|\,\{N_n(s)\}_{s\leqslant t},\hat{Z}_n)\\&=N^0_{An}(t-c)Q_{N_n(t)-n,n}\mathbf{1}_{[c,\infty)}(t)\\&=\{2n-N_n(t)+Z_n(t-c)\}Q_{N_n(t)-n,n}\mathbf{1}_{[c,\infty)}(t).\end{aligned}\tag{3.10}$$

The process stops at time T_n given by (3.8) and (3.9) or equivalently by

$$T_n=\inf\left\{t:\ N_n(t)=n+\inf\left\{m:\ \sum_{i=1}^{m}(Z_{in}-1)=-n\right\}\wedge n_{sn}\right\},$$

which is clearly a stopping time with respect to the σ-algebra generated by $\{N_n(s)\}_{s\leqslant t}$ and $\hat{Z}_n$.

Let $A_{zn}(t)$ be the integrated intensity process of $N_n(t)-n$ given $\hat{Z}_n$, i.e.

$$A_{zn}(t)=\int_0^t\Lambda_{zn}(s)\,\mathrm{d}s.\tag{3.11}$$

Define the inverse process A^{-1}_{zn} by

$$A^{-1}_{zn}(u)=\begin{cases}\inf\{t\geqslant c:\ A_{zn}(t)\geqslant u\}, & 0\leqslant u\leqslant A_{zn}(T_n),\\ \infty, & u>A_{zn}(T_n),\end{cases}\tag{3.12}$$

and the process N^*_n by

$$N^*_n(u)=\begin{cases}N_n(A^{-1}_{zn}(u))-n, & 0\leqslant u\leqslant A_{zn}(T_n),\\ \Pi_n(u-A_{zn}(T_n))+N_n(T_n)-n, & u>A_{zn}(T_n),\end{cases}\tag{3.13}$$

where Π_n is a Poisson process with intensity 1 defined on $[0,\infty)$, independent of N_n and Λ_{zn}. Then given $\hat{Z}_n$, N^*_n is a Poisson process with intensity 1 on $[0,\infty)$. It follows that unconditionally N^*_n is also a unit Poisson process which is independent of $\hat{Z}_n$. Note that the stopping time T_n, however, is a function of $\hat{Z}_n$. For a detailed discussion on counting processes and other concepts mentioned above see for instance Jacobsen (1982).

In the following sections we investigate the behaviour of N_n for n tending to infinity. We note that some of the results derived can be arrived at by using known theory for population processes, which can be found in Kurtz (1981, 1983) or Ethier and Kurtz (1986). To apply the theory directly to our counting process N_n as defined above, we need an explicit expression in terms of $\{N_n(s)\}_{s\leq t}$ for its unconditional intensity process $\Lambda_n(t) = E_Z(\Lambda_{Z_n}(t))$. However, such an expression is not available in a manageable form. One way to overcome this problem is to consider a trivariate counting process $\tilde{N}_n(t)$ which has as components the number of stimuli before or at time t resulting in two A-cells, one A-cell and one B-cell, and two B-cells, respectively, as is done in De Gunst (1989). The intensity process for this trivariate process is easily available. Apart from a constant time shift, the process N_n is the sum of the components of $\tilde{N}_n$, and N_n inherits some of the asymptotic properties of $\tilde{N}_n$.

The approach adopted in this paper is to consider the process N_n conditionally on the sequence $\hat{Z}_n = (Z_{1n}, Z_{2n}, \ldots)$. The conditional intensity process of N_n has a structure very similar to the unconditional intensity of the trivariate process. This is why the uniform convergence of $n^{-1}(N_n - n)$ can be shown here in essentially the same manner as that of the trivariate process. Also the proofs of the central limit theorem in both approaches are similar. However, the representation of the limit distribution in terms of two independent Wiener processes as derived below ((5.8) and (5.9)), and its covariance structure are almost impossible to guess without going into the special structure of the underlying process as is done in Lemma 5.2. To obtain further asymptotic properties of N_n, the special character of the process must play a crucial role. For instance, the uniform convergence of $n^{-1}(N_n - n)$ on $[0, \infty)$ appears to depend heavily on the boundedness of $n^{-1}(N_n - n)$.

4. The asymptotic behaviour of N_n

In this section we prove several probability inequalities which enable us to derive an exponential bound for the probability that $\sup_{t\geq 0} |n^{-1}(N_n(t) - n) - X(t)|$ exceeds some value x. Here X is a non-random differentiable function of t, which will be defined later. Uniform convergence of $n^{-1}(N_n - n)$ to X and the rate of convergence are then easily established.

Let us first state two more general results, which will be needed below. For their proofs we refer to De Gunst (1989).

Lemma 4.1. *If $\Pi(t)$, $t \geq 0$ is a unit Poisson process, then*

$$\mathcal{P}\Big(\sup_{\{t:\,\Pi(t)\leq a\}} |\Pi(t) - t| \geq x\Big) \leq \tfrac{11}{3}\exp\{-x^2/(32a)\}, \quad 0 \leq x \leq 2a. \qquad \square \qquad (4.1)$$

Lemma 4.2. *If* $X_1, X_2, \ldots$ *are independent, bounded random variables,* $0 \leqslant X_j \leqslant a$, $j = 1, 2, \ldots$, *then for all* $M \in \mathbb{N}$ *and every* $x \geqslant 0$,

$$\mathcal{P}\left(\max_{1 \leqslant m \leqslant M} \left| \sum_{j=1}^{m} (X_j - EX_j) \right| \geqslant x \right) \leqslant 4 \exp\{-2x^2/(9a^2M)\}. \quad \square \tag{4.2}$$

From now on we write

$$X_n(t) = n^{-1}(N_n(t) - n).$$

We have:

Lemma 4.3. *For* $n = 1, 2, \ldots$ *and* $0 \leqslant x \leqslant 2b_{\mathrm{s}}$,

$$\mathcal{P}\left(\sup_{t \geqslant c} |X_n(t) - n^{-1}\Lambda_{zn}(t)| \geqslant x \right) \leqslant \tfrac{11}{3} \exp\{-x^2 n/(32 b_{\mathrm{s}})\}. \tag{4.3}$$

Proof. Since $X_n(t) \leqslant n^{-1} n_{\mathrm{s}n} = b_{\mathrm{s}}$, we see that

$$\sup_{t \geqslant c} |X_n(t) - n^{-1}\Lambda_{zn}(t)| \leqslant \sup_{\{u:\ N_n^*(u) \leqslant nb_{\mathrm{s}}\}} n^{-1}|N_n^*(u) - u|,$$

where N_n^* is the unit Poisson process defined in (3.13). The lemma now follows from Lemma 4.1. $\square$

Define

$$\Lambda_{pn}(t) = \left\{ 2n - N_n(t) + 2 \sum_{i=1}^{nX_n(t-c)} P_{in} \right\} Q_{nX_n(t),n} \mathbf{1}_{[c,\infty)}(t), \tag{4.4}$$

and note that Λ_{pn} is obtained from Λ_{zn} in (3.10) by replacing the Z_{in} by their expected values $2P_{in}$, so that:

Lemma 4.4. *For* $n = 1, 2, \ldots$ *and* $x \geqslant 0$,

$$\mathcal{P}\left(\max_{m} n^{-1} \left| \sum_{i=1}^{m} (Z_{in} - 2P_{in}) \right| \geqslant x \right) \leqslant 4 \exp\{-x^2 n/(18 b_{\mathrm{h}})\}, \tag{4.5}$$

$$\mathcal{P}\left(\sup_{t \geqslant c} n^{-1} |\Lambda_{zn}(t) - \Lambda_{pn}(t)| \geqslant x \right) \leqslant 4 \exp\{-d^2(1 + b_{\mathrm{s}}^{-1} k_{\mathrm{s}} y_{\mathrm{s}})^2 x^2 n/(18 b_{\mathrm{h}})\}. \tag{4.6}$$

Proof. $Z_{1n}, Z_{2n}, \ldots$ are independent and Z_{in} has a binomial distribution with parameters 2 and P_{in}, $i = 1, 2, \ldots$; for $i > n_{\mathrm{h}n}$, $P_{in} = 0$ and $Z_{in} = 0$ a.s. Since $n_{\mathrm{h}n} = nb_{\mathrm{h}}$, we may invoke Lemma 4.2 to obtain (4.5).

Furthermore, we see from definitions (3.10), (4.4) and (3.3) that for $t \geqslant c$,

$$\begin{aligned} |\Lambda_{zn}(t) - \Lambda_{pn}(t)| &= \left| \sum_{i=1}^{nX_n(t-c)} (Z_{in} - 2P_{in}) Q_{nX_n(t),n} \right| \\ &\leqslant (d(1 + b_{\mathrm{s}}^{-1} k_{\mathrm{s}} y_{\mathrm{s}}))^{-1} \max_{1 \leqslant m \leqslant n_{\mathrm{h}n}} \left| \sum_{i=1}^{m} (Z_{in} - 2P_{in}) \right|. \end{aligned}$$

Hence we get (4.6) from (4.5). $\square$

Let Q, P and $\bar{P}:[0,\infty)\to[0,\infty)$ be defined by

$$Q(u)=\frac{b_{\mathrm{s}}-(u\wedge b_{\mathrm{s}})}{d(b_{\mathrm{s}}+k_{\mathrm{s}}y_{\mathrm{s}}-(u\wedge b_{\mathrm{s}}))}, \tag{4.7a}$$

$$P(u)=\frac{b_{\mathrm{h}}-(u\wedge b_{\mathrm{h}})}{b_{\mathrm{h}}+k_{\mathrm{h}}y_{\mathrm{h}}-(u\wedge b_{\mathrm{h}})}, \tag{4.7b}$$

$$\bar{P}(u)=\int_0^u P(s)\,\mathrm{d}s=(u\wedge b_{\mathrm{h}})-k_{\mathrm{h}}y_{\mathrm{h}}\log\left(\frac{b_{\mathrm{h}}+k_{\mathrm{h}}y_{\mathrm{h}}}{b_{\mathrm{h}}+k_{\mathrm{h}}y_{\mathrm{h}}-(u\wedge b_h)}\right), \tag{4.7c}$$

and define $F:[0,\infty)\times D[0,\infty)\to\mathbb{R}$ by

$$F(t,x)=\{1-x(t)+2\bar{P}(x(t-c))\}Q(x(t))\mathbf{1}_{[c,\infty)}(t), \tag{4.8}$$

where $D[0,\infty)$ is the space of right-continuous, $\mathbb{R}$-valued functions on $[0,\infty)$ with left-hand limits everywhere. Obviously, for $t\geqslant 0$ and continuous functions x and y on $[0,\infty)$, we have

$$\sup_{s\leqslant t}|F(s,x)|\leqslant\frac{1}{d}\left(1+\sup_{s\leqslant t}|x(s)|\right) \tag{4.9}$$

and since the partial derivatives of F with respect to $x(t)$ and $x(t-c)$ are bounded,

$$\begin{aligned}|F(t,x)-F(t,y)|&\leqslant C|x(t)-y(t)|+C'|x(t-c)-y(t-c)|\\&\leqslant C^*\sup_{s\leqslant t}|x(s)-y(s)|\end{aligned} \tag{4.10}$$

for positive constants C, C' an C^*. Note that

$$Q_{N_n(t)-n,n}=Q_{nX_n(t),n}=Q(X_n(t)), \tag{4.11}$$

$$P_{N_n(t)-n,n}=P_{nX_n(t),n}=P(X_n(t))+\mathrm{O}(1/n), \tag{4.12}$$

so that $F(t,X_n)$ may be expected to approximate $n^{-1}\Lambda_{pn}(t)$. In fact, we have:

Lemma 4.5. *For* $n=1,2,\ldots,$

$$\max_m\left|\sum_{i=1}^m P_{in}-n\bar{P}(m/n)\right|\leqslant P(0)=(1+b_{\mathrm{h}}^{-1}k_{\mathrm{h}}y_{\mathrm{h}})^{-1}, \tag{4.13}$$

$$\sup_{t\geqslant c}|n^{-1}\Lambda_{pn}(t)-F(t,X_n)|\leqslant 2(d(1+b_{\mathrm{s}}^{-1}k_{\mathrm{s}}y_{\mathrm{s}})n)^{-1}. \tag{4.14}$$

Proof. From (4.7) it is easy to see that

$$0\leqslant\sum_{i=1}^m P_{in}-n\bar{P}(m/n)\leqslant\sum_{i=1}^{m\wedge n_{\mathrm{h}n}}(P((i-1)/n)-P(i/n))\leqslant P(0),$$

so that (4.13) follows immediately. Next, we recall that (cf. (3.3))

$$\sup_{t\geqslant c}Q(X_n(t))\leqslant(d(1+b_{\mathrm{s}}^{-1}k_{\mathrm{s}}y_{\mathrm{s}}))^{-1}$$

and in view of (4.4), (4.8) and (4.13) we conclude that for $t \geq c$,

$$\begin{aligned}|n^{-1}\Lambda_{pn}(t) - F(t, X_n)| &= 2\left| n^{-1} \sum_{i=1}^{nX_n(t-c)} P_{in} - \bar{P}(X_n(t-c))\right| Q(X_n(t)) \\ &\leq 2(d(1+b_{\mathrm{h}}^{-1}k_{\mathrm{h}}y_{\mathrm{h}})(1+b_{\mathrm{s}}^{-1}k_{\mathrm{s}}y_{\mathrm{s}})n)^{-1} \\ &\leq 2(d(1+b_{\mathrm{s}}^{-1}k_{\mathrm{s}}y_{\mathrm{s}})n)^{-1}. \qquad \square\end{aligned}$$

Consider the system

$$x(t) = \begin{cases} \int_c^t F(s, x)\, \mathrm{d}s, & t \geq c, \\ 0, & t < c, \end{cases} \tag{4.15}$$

with $x \in D[0, \infty)$ and F as defined in (4.8). Then it is easy to prove (see De Gunst, 1989) that (4.9) and (4.10) are the conditions for a unique solution $X \in D[0, \infty)$ of (4.15) to exist, and that X is continuous on $(0, \infty)$. Since $X'(t) = F(t, X)$ on (c, ∞) and $F(t, X) = 0$ if $X(t) \geq b_{\mathrm{s}}$, we see that X is bounded on $(0, \infty)$ and differentiable on (c, ∞) with a continuous and bounded derivative.

Having made these preparations, we can prove the following result.

Theorem 4.1. *Let X be the unique solution of* (4.15) *and let $M \geq 0$. Then there exist positive numbers A and α depending on M such that for $n = 1, 2, \ldots$ and $x \geq 0$,*

$$\mathscr{P}\left(\sup_{t \leq M} |X_n(t) - X(t)| \geq x\right) \leq A \exp\{-\alpha x^2 n\}. \tag{4.16}$$

Proof. Fix $M \geq c$, let C^* be as in (4.10), $C = 6b_{\mathrm{s}} \exp\{C^*(M-c)\}$ and $6(M-c)\exp\{C^*(M-c)\}(d(1+b_{\mathrm{s}}^{-1}k_{\mathrm{s}}y_{\mathrm{s}})n)^{-1} \leq x \leq C$. Note that $C \geq 6b_{\mathrm{s}}$. Also, for every $n = 1, 2, \ldots$, it follows from Lemmas 4.3, 4.4, 4.5 and inequality (4.10) that, except on a set of probability not larger than

$$\begin{aligned}&\tfrac{11}{3}\exp\{-x^2 n/[288 b_{\mathrm{s}} \exp\{2C^*(M-c)\}]\} \\ &\quad + 4\exp\{-d^2(1+a_{\mathrm{s}})^2 x^2 n/[162 b_{\mathrm{h}}(M-c)^2 \exp\{2C^*(M-c)\}]\},\end{aligned}$$

the following holds for all $c \leq t \leq M$:

$$\begin{aligned}|X_n(t) - X(t)| &\leq |X_n(t) - n^{-1}\Lambda_{zn}(t)| + n^{-1}\int_c^t |\Lambda_{zn}(s) - \Lambda_{pn}(s)|\, \mathrm{d}s \\ &\quad + \int_c^t |n^{-1}\Lambda_{pn}(s) - F(s, X_n)|\, \mathrm{d}s + \int_c^t |F(s, X_n) - F(s, X)|\, \mathrm{d}s \\ &\leq 3\frac{x}{3}\exp\{-C^*(M-c)\} + C^* \int_c^t \sup_{\tau \leq s} |X_n(\tau) - X(\tau)|\, \mathrm{d}s \\ &\leq x\exp\{-C^*(M-c)\}\exp\{C^*(t-c)\} \leq x,\end{aligned}$$

where the third inequality follows from a version of Gronwall's lemma (see De Gunst, 1989).

Obviously, A can be chosen sufficiently large for (4.16) to hold for $0 \leq x < 6(M-c)\exp\{C^*(M-c)\}(d(1+b_s^{-1}k_s y_s)n)^{-1}$ and since $X_n(t) \leq b_s$ and $X(t) \leq b_s$ for all $t \geq c$, (4.16) is trivially satisfied for $x > C \geq b_s$. By definition $X_n(t) = X(t) = 0$ for $0 \leq t < c$ and all n, and the theorem follows. □

In order to obtain a probability inequality like (4.16) for the supremum over all t, we first investigate what happens when $t \to \infty$. Since $X'(t)$ can be proved to be positive for $t > c$ (De Gunst, 1989) and $X''(t)$ is easily seen to be bounded for $t > 2c$, it follows that $X'(t)$ tends to zero as t tends to infinity. Writing $X(\infty)$ for $\lim_{t\to\infty} X(t)$, we find that $\lim_{t\to\infty} X'(t) = (1 - X(\infty) + 2\bar{P}(X(\infty)))Q(X(\infty)) = 0$, and since for $t > c$, $(1 - X(t) + 2\bar{P}(X(t)))Q(X(t)) \geq X'(t) > 0$, we see

$$\lim_{t\to\infty} X(t) = \inf\{z \in \mathbb{R}: 2\bar{P}(z) - z \leq -1\} \wedge b_s. \tag{4.17}$$

Recall (cf. (3.9)) that

$$\lim_{t\to\infty} X_n(t) = n^{-1} \inf\left\{m: \sum_{i=1}^{m} (Z_{in} - 1) = -n\right\} \wedge b_s, \tag{4.18}$$

so that (4.5) and (4.13) suggest that $\lim_{t\to\infty} X_n(t)$ approaches $X(\infty)$ as n tends to infinity. Indeed, a comparison of both the right-hand side of (4.17) and (4.18) with $n^{-1}\inf\{m: \sum_{i=1}^{m} (2P_{in} - 1) \leq -n\} \wedge b_s$ yields:

Lemma 4.6. *If X is the solution of* (4.15) *on* $[0,\infty)$, *then there exist positive numbers A and α such that for $n = 1, 2, \ldots$ and $x \geq 0$,*

$$\mathcal{P}\left(\left|\lim_{t\to\infty} X_n(t) - \lim_{t\to\infty} X(t)\right| \geq x\right) \leq A\exp\{-\alpha x^2 n\}. \tag{4.19}$$

Proof. Clearly, for any A, α and B such that $A > \exp\{4\alpha B^2\}$ the lemma holds for $0 \leq x < 2Bn^{-1}$. Hence, in view of (4.17) and (4.18) it suffices to prove that there exist positive constants A, α and B such that for $n = 1, 2, \ldots$ and $x \geq 0$,

$$\text{(i)}\quad \mathcal{P}\left(n^{-1}\left|\left(n + \inf\left\{m: \sum_{i=1}^{m} (Z_{in} - 1) = -n\right\} \wedge n_{sn}\right) - \left(n + \inf\left\{m: \sum_{i=1}^{m} (2P_{in} - 1) \leq -n\right\} \wedge n_{sn}\right)\right| \geq x\right) \leq A\exp\{-\alpha x^2 n\},$$

$$\text{(ii)}\quad \left|n^{-1}\left(n + \inf\left\{m: \sum_{i=1}^{m} (2P_{in} - 1) \leq -n\right\} \wedge n_{sn}\right) - (1 + \inf\{z \in \mathbb{R}: 2\bar{P}(z) - z \leq -1\} \wedge b_s)\right| \leq Bn^{-1}.$$

To prove (i) first note that $|a \wedge c - b \wedge c| \leq |a - b|$. Moreover, since (i) is trivially true for $0 \leq nx \leq 1$ for an appropriate choice of α, we may assume without loss of generality that nx is a positive integer. Similarly, we allow ourselves in the proof of (i) to bound the probability of strict inequality. Hence we have to prove that for some $\alpha>0$, $n=1,2,\ldots$, $x=2/n,\ 3/n,\ldots$,

$$\mathcal{P}\left(n^{-1}\left|\inf\left\{m: \sum_{i=1}^{m}(Z_{in}-1)=-n\right\}-\inf\left\{m: \sum_{i=1}^{m}(2P_{in}-1)\leq -n\right\}\right|>x\right)$$
$$\leq A\exp\{-\alpha x^2 n\}.$$

Let $m_0=\inf\{m: \sum_{i=1}^{m}(2P_{in}-1)\leq -n\}$. Then

$$n^{-1}\sum_{i=1}^{m_0}(2P_{in}-1)\leq -1,\qquad n^{-1}\sum_{i=1}^{m_0-1}(2P_{in}-1)>-1.$$

Since P_{in} decreases in i and $m_0\leq 2n_{\mathrm{h}n}+n=\mathrm{O}(n)$, for some $\gamma>0$,

$$2P_{m_0 n}-1\leq -n/m_0\leq -\gamma<0,$$
$$n^{-1}\sum_{i=1}^{m_0+nx}(2P_{in}-1)\leq -1-\gamma x \quad \text{for } x=\frac{2}{n},\frac{3}{n},\ldots.$$

Using this and (4.5), we get

$$\mathcal{P}\left(n^{-1}\inf\left\{m: \sum_{i=1}^{m}(Z_{in}-1)=-n\right\}>n^{-1}m_0+x\right)$$
$$=\mathcal{P}\left(n^{-1}\min_{k\leq m_0+nx}\sum_{i=1}^{k}(Z_{in}-1)>-1\right)$$
$$\leq\mathcal{P}\left(n^{-1}\sum_{i=1}^{m_0+nx}(Z_{in}-1)>-1\right)$$
$$=\mathcal{P}\left(n^{-1}\sum_{i=1}^{m_0+nx}(Z_{in}-2P_{in})>-1+n^{-1}\sum_{i=1}^{m_0+nx}(1-2P_{in})\right)$$
$$\leq\mathcal{P}\left(n^{-1}\sum_{i=1}^{m_0+nx}(Z_{in}-2P_{in})\geq\gamma x\right)$$
$$\leq 4\exp\{-\gamma^2x^2n/(18b_{\mathrm{h}})\}.$$

On the other hand, because $\sum_{i=1}^{m}(2P_{in}-1)$ is a concave function of m which vanishes for $m=0$, $m^{-1}\sum_{i=1}^{m}(2P_{in}-1)$ is a decreasing one. Since

$$-n<\sum_{i=1}^{m_0-1}(2P_{in}-1)\leq -n+1\leq 0,$$

if follows that for some positive δ and $1\leq k\leq m_0-1-nx$, $x=2/n,\ 3/n,\ldots$,

$$n^{-1}\sum_{i=1}^{k}(2P_{in}-1)\geq k(m_0-1)^{-1}n^{-1}\sum_{i=1}^{m_0-1}(2P_{in}-1)$$
$$\geq(m_0-1-nx)(m_0-1)^{-1}n^{-1}\sum_{i=1}^{m_0-1}(2P_{in}-1)$$
$$\geq-(1-nx(2n_{\mathrm{h}n}+n)^{-1})\geq -1+\delta x,$$

provided $m_0 - 1 - nx \geqslant 0$. But this and (4.5) imply

$$\begin{aligned}
&\mathscr{P}\left(n^{-1}\inf\left\{m\colon \sum_{i=1}^{m}(Z_{in}-1)=-n\right\}<n^{-1}m_0-x\right)\\
&\quad=\mathscr{P}\left(n^{-1}\min_{k\leqslant m_0-nx-1}\sum_{i=1}^{k}(Z_{in}-1)\leqslant -1\right)\\
&\quad\leqslant\mathscr{P}\left(n^{-1}\min_{k\leqslant m_0-nx-1}\sum_{i=1}^{k}(Z_{in}-2P_{in})\leqslant -\delta x\right)\\
&\quad\leqslant 4\exp\{-\delta^2x^2n/(18b_{\mathrm{h}})\}.
\end{aligned}$$

For $m_0 - 1 - nx \leqslant 0$ the first probability equals zero and the inequality is trivially correct. This proves (i).

The proof of (ii) is essentially similar, except that (4.13) is used instead of (4.5). □

Having proved the probability bounds (4.16) and (4.19) for bounded t and $t \to \infty$, respectively, we can now derive a similar result which holds uniformly for all $t \geqslant 0$.

For convenience we introduce a shorter notation for the quantities associated with the newly born A-cells. Let Y_n and Y be defined by

$$Y_n(t) = n^{-1}Z_n(t) = n^{-1}\sum_{i=1}^{nX_n(t)} Z_{in}, \tag{4.20a}$$

$$Y(t) = 2\bar{P}(X(t)) = 2\int_0^{X(t)} P(u)\,\mathrm{d}u = 2\int_0^t P(X(s))X'(s)\,\mathrm{d}s. \tag{4.20b}$$

We also write

$$X^0_{\mathrm{A}n}(t-c) = 1 - X_n(t) + Y_n(t-c) \geqslant 0, \qquad \alpha_n = n^{-1}A_{zn}, \tag{4.21a}$$

$$X^0_{\mathrm{A}}(t-c) = 1 - X(t) + Y(t-c) \geqslant 0, \tag{4.21b}$$

$$p = P(X(\infty)) \geqslant 0, \quad q = Q(X(\infty)) \geqslant 0, \quad q' = Q'(X(\infty)) < 0. \tag{4.21c}$$

Recall that for $t > c$, we have (cf. (4.8), (4.15), (3.10), (3.11) and (4.11))

$$X'(t) = F(t, X) = X^0_{\mathrm{A}}(t-c)Q(X(t)), \tag{4.22a}$$

$$\alpha'_n(t) = n^{-1}A_{zn}(t) = X^0_{\mathrm{A}n}(t-c)Q(X_n(t)), \tag{4.22b}$$

where α'_n denotes the derivative of α_n except on a countable set of jumps of X_n. Also, note that for $0 \leqslant u \leqslant X(\infty)$,

$$0 \geqslant P'(u) > -(k_{\mathrm{h}}y_{\mathrm{h}})^{-1}, \qquad 0 \geqslant P''(u) > -2(k_{\mathrm{h}}y_{\mathrm{h}})^{-2}, \tag{4.23a}$$

$$0 > Q'(u) > -(dk_{\mathrm{s}}y_{\mathrm{s}})^{-1}, \qquad 0 \geqslant Q''(u) > -2d^{-1}(k_{\mathrm{s}}y_{\mathrm{s}})^{-2}. \tag{4.23b}$$

We need the following lemma.

Lemma 4.7. *Suppose that* $\varepsilon > 0$ *and define*

$$t_0 = \inf\{t \geqslant c\colon X(\infty) - X(t-c) \leqslant \tfrac{1}{2}\varepsilon k_{\mathrm{h}} y_{\mathrm{h}}\}. \tag{4.24}$$

Then there exist positive numbers A_0, A_1, α_0 *and* α_1 *depending on* ε, *such that for* $n = 1, 2, \ldots$ *and* $x \geqslant 0$,

$$\mathscr{P}\Big(\sup_{t \geqslant t_0} |X(\infty) - X_n(t-c)| \geqslant \varepsilon k_{\mathrm{h}} y_{\mathrm{h}}\Big) \leqslant A_0 \exp\{-\alpha_0 n\}, \tag{4.25}$$

$$\begin{aligned} &\mathscr{P}\Big(\sup_{t \geqslant t_0} [|Y_n(t-c) - Y(t-c)| - 2(p+\varepsilon)|X_n(t-c) - X(t-c)|] \geqslant x\Big) \\ &\quad \leqslant A_1 \exp\{-\alpha_1 x^2 n\}. \end{aligned} \tag{4.26}$$

Proof. Since X_n is monotone, we have for $t \geqslant t_0$,

$$\begin{aligned} |X(\infty) - X_n(t-c)| &\leqslant |X(\infty) - X_n(T_n)| \vee |X(\infty) - X_n(t_0 - c)| \\ &\leqslant |X_n(T_n) - X(\infty)| + |X_n(t_0-c) - X(t_0-c)| + \tfrac{1}{2}\varepsilon k_{\mathrm{h}} y_{\mathrm{h}} \end{aligned}$$

and in view of Lemma 4.6 and Theorem 4.1,

$$\begin{aligned} &\mathscr{P}\Big(\sup_{t \geqslant t_0} |X(\infty) - X_n(t-c)| \geqslant \varepsilon k_{\mathrm{h}} y_{\mathrm{h}}\Big) \\ &\quad \leqslant \mathscr{P}(|X_n(T_n) - X(\infty)| \geqslant \tfrac{1}{4}\varepsilon k_{\mathrm{h}} y_{\mathrm{h}}) + \mathscr{P}(|X_n(t_0-c) - X(t_0-c)| \geqslant \tfrac{1}{4}\varepsilon k_{\mathrm{h}} y_{\mathrm{h}}) \\ &\quad \leqslant A_0 \exp\{-\alpha_0 n\}. \end{aligned}$$

for positive A_0, α_0 depending on ε. This proves (4.25).

For $0 \leqslant x \leqslant 8/n$, (4.26) trivially holds for an appropriate choice of A_1 and α_1. Since $Y_n(t) \leqslant 2X_n(t) \leqslant 2b_{\mathrm{s}}$ and $Y(t) \leqslant 2X(t) \leqslant 2b_{\mathrm{s}}$ for all t, (4.26) is also satisfied for $x > 4b_{\mathrm{s}}$. Assume therefore that $8/n < x \leqslant 4b_{\mathrm{s}}$. By Lemma 4.4,

$$\begin{aligned} &\mathscr{P}\Bigg(\sup_{t \geqslant c} n^{-1}\Bigg|\sum_{i=1}^{nX_n(t-c)} (Z_{in} - 2P_{in})\Bigg| \geqslant \tfrac{1}{2}x - \frac{2}{n}\Bigg) \\ &\quad \leqslant \mathscr{P}\Bigg(\max_m n^{-1}\Bigg|\sum_{i=1}^{m} (Z_{in} - 2P_{in})\Bigg| \geqslant \tfrac{1}{4}x\Bigg) \leqslant A_2 \exp\{-\alpha_2 x^2 n\} \end{aligned}$$

for positive constants A_2 and α_2. From (4.13) we know that with probability 1,

$$2\Bigg|n^{-1} \sum_{i=1}^{nX_n(t-c)} P_{in} - \bar{P}(X_n(t-c))\Bigg| \leqslant \frac{2}{n} \quad \text{for all } t \geqslant c.$$

Finally, since $P(X(\infty)) = p$ and $|P'(u)| < (k_{\mathrm{h}} y_{\mathrm{h}})^{-1}$ for all u, we find that if $t \geqslant t_0$ and $|X(\infty) - X_n(t-c)| < \varepsilon k_{\mathrm{h}} y_{\mathrm{h}}$,

$$\begin{aligned} &|2\bar{P}(X_n(t-c)) - Y(t-c)| \\ &\quad = 2|\bar{P}(X_n(t-c)) - \bar{P}(X(t-c))| \\ &\quad \leqslant 2[p + (k_{\mathrm{h}} y_{\mathrm{h}})^{-1}\{(X(\infty) - X(t-c)) \vee |X(\infty) - X_n(t-c)|\}] \\ &\qquad \times |X_n(t-c) - X(t-c)| \\ &\quad \leqslant 2(p+\varepsilon)|X_n(t-c) - X(t-c)|. \end{aligned}$$

Hence, in view of (4.24) and (4.25), and since $\frac{1}{2}x > 4/n$,

$$
\begin{aligned}
&\mathscr{P}\Big(\sup_{t\geq t_0}[|Y_n(t-c)-Y(t-c)|-2(p+\varepsilon)|X_n(t-c)-X(t-c)|]>x\Big)\\
&\quad\leq \mathscr{P}\Bigg(\sup_{t\geq t_0}\Bigg[n^{-1}\Bigg|\sum_{i=1}^{nX_n(t-c)}(Z_{in}-2P_{in})\Bigg|\\
&\qquad\qquad +2\Bigg|n^{-1}\sum_{i=1}^{nX_n(t-c)}P_{in}-\bar{P}(X_n(t-c))\Bigg|\Bigg]\geq\tfrac{1}{2}x\Bigg)\\
&\qquad +\mathscr{P}\Big(\sup_{t\geq t_0}[|2\bar{P}(X_n(t-c))-Y(t-c)|\\
&\qquad\qquad -2(p+\varepsilon)|X_n(t-c)-X(t-c)|)]\geq\frac{4}{n}\Big)\\
&\quad\leq A_2\exp\{-\alpha_2x^2n\}+\mathscr{P}\Big(\sup_{t\geq t_0}|X(\infty)-X_n(t-c)|\geq\varepsilon k_{\mathrm{h}}y_{\mathrm{h}}\Big)\\
&\quad\leq A_2\exp\{-\alpha_2x^2n\}+A_0\exp\{-\alpha_0n\}\\
&\quad\leq A_2\exp\{-\alpha_2x^2n\}+A_0\exp\{-\alpha_0x^2n/(4b_{\mathrm{s}})^2\}\leq A_1\exp\{-\alpha_1x^2n\}
\end{aligned}
$$

for appropriate positive A_1 and α_1 depending on ε. □

Now we are ready to prove the main result of this section.

Theorem 4.2. *Let X be the solution of* (4.15). *Then there exist positive numbers A and α such that for $n=1,2,\ldots$ and $x\geq 0$,*

$$
\mathscr{P}\Big(\sup_{t\geq 0}|X_n(t)-X(t)|\geq x\Big)\leq A\exp\{-\alpha x^2n\}. \tag{4.27}
$$

Proof. Let T_n be the time X_n stops as defined by (3.8). Since X is monotone, we have

$$
\sup_{t\geq 0}|X_n(t)-X(t)|\leq\sup_{0\leq t\leq T_n}|X_n(t)-X(t)|+\Big|X_n(T_n)-\lim_{t\to\infty}X(t)\Big|.
$$

Therefore, in view of Lemma 4.6 it suffices to show that

$$
\mathscr{P}\Big(\sup_{0\leq t\leq T_n}|X_n(t)-X(t)|\geq x\Big)\leq A\exp\{-\alpha x^2n\}. \tag{4.28}
$$

The proof consists of two parts. First suppose that the following assumption holds.

(A) There exist $t_0\geq c$, positive constants $\tilde{A}$, $\tilde{\alpha}$ and $\tilde{C}$, and non-negative functions M_n, U_n, V_n with

$$
U_n(t)+V_n(t)-M_n(t)<0,\quad t_0<t\leq T_n,
$$

such that for $n=1,2,\ldots,$ and $0\leq x\leq\tilde{C}$,

$$
\mathscr{P}\Big(\sup_{t\leq t_0}|\alpha_n(t)-X(t)|\geq x\Big)\leq\tilde{A}\exp\{-\tilde{\alpha}x^2n\}, \tag{4.29}
$$

$$\alpha_n'(t)-X'(t)=-M_n(t)\{\alpha_n(t)-X(t)\}+R_n(t), \quad t_0<t\leqslant T_n, \tag{4.30a}$$

with R_n such that

$$\mathscr{P}(|R_n(t)|\geqslant U_n(t)x+V_n(t)|\alpha_n(t-c)-X(t-c)| \text{ for some } t\in(t_0,T_n])\leqslant \tilde{A}\exp\{-\tilde{\alpha}x^2n\}. \tag{4.30b}$$

Fix n and $x\in(0,\tilde{C}]$. Suppose that there exists $t_1\in(t_0,T_n]$ such that $|\alpha_n(t)-X(t)|<x$ for $t<t_1$ and $|\alpha_n(t_1)-X(t_1)|=x$, and that

$$|R_n(t_1)|<U_n(t_1)x+V_n(t_1)|\alpha_n(t_1-c)-X(t_1-c)|.$$

Then

$$\begin{aligned}|R_n(t_1)|&<U_n(t_1)x+V_n(t_1)|\alpha_n(t_1-c)-X(t_1-c)|\\&\leqslant(U_n(t_1)+V_n(t_1))|\alpha_n(t_1)-X(t_1)|.\end{aligned}$$

Hence

$$(\alpha_n(t_1)-X(t_1))^{-1}(\alpha_n'(t_1)-X'(t_1))\leqslant -M_n(t_1)+U_n(t_1)+V_n(t_1)<0$$

which contradicts the assumption that t_1 is the first time when $\alpha_n(t)-X(t)=\pm x$. It follows from (4.29) and (4.30) that for $x\in(0,\tilde{C}]$,

$$\begin{aligned}&\mathscr{P}\Big(\sup_{0\leqslant t\leqslant T_n}|\alpha_n(t)-X(t)|\geqslant x\Big)\\&\leqslant\mathscr{P}\Big(\sup_{t\leqslant t_0}|\alpha_n(t)-X(t)|\geqslant x\Big)\\&\quad+\mathscr{P}(|R_n(t)|\geqslant U_n(t)x+V_n(t)|\alpha_n(t-c)-X(t-c)|\\&\qquad\qquad \text{for some } t\in(t_0,T_n])\leqslant 2\tilde{A}\exp\{-\tilde{\alpha}x^2n\}.\end{aligned}$$

Combining this with Lemma 4.3 we obtain (4.28) for $x\in(0,C]$ for some positive C. If $A\geqslant 1$, then (4.28) continues to hold for $x=0$. Since $|X_n(t)-X(t)|\leqslant 2b_s$ for all t, we see that for $x>C$, the left-hand side of (4.28) is bounded by $A\exp\{-\alpha x^2nC^2/(2b_s)^2\}$. Thus the theorem is proved under assumption (A).

We now turn to the proof of assumption (A). We shall distinguish two cases.

Case (i). $p<\frac{1}{2}$. Choose $\varepsilon=\frac{1}{10}(1-2p)$ and let t_0 be as in (4.24). From Lemma 4.3, Theorem 4.1 and Lemma 4.7 it follows that there exist positive numbers $\tilde{A}$, $\tilde{\alpha}$ and $\tilde{C}$ depending on ε, such that for $n=1,2,\ldots$ and $0\leqslant x\leqslant\tilde{C}$,

$$\mathscr{P}\Big(\sup_{t\geqslant 0}|\alpha_n(t)-X_n(t)|\geqslant\varepsilon x\Big)\leqslant\tfrac{1}{3}\tilde{A}\exp\{-\tilde{\alpha}x^2n\}, \tag{4.31a}$$

$$\mathscr{P}\Big(\sup_{t\leqslant t_0}|\alpha_n(t)-X(t)|\geqslant x\Big)\leqslant\tfrac{1}{3}\tilde{A}\exp\{-\tilde{\alpha}x^2n\}, \tag{4.31b}$$

$$\begin{aligned}&\mathscr{P}\Big(\sup_{t\geqslant t_0}[|Y_n(t-c)-Y(t-c)|-2(p+\varepsilon)|X_n(t-c)-X(t-c)|]\geqslant\varepsilon x\Big)\\&\leqslant\tfrac{1}{3}\tilde{A}\exp\{-\tilde{\alpha}x^2n\}.\end{aligned} \tag{4.31c}$$

For $c<t\leqslant T_n$ we have $\alpha_n'(t)=X^0_{\mathrm{A}n}(t-c)Q(X_n(t))$ (cf. (4.22)). Hence, except on a set of probability not larger than $\tilde{A}\exp\{-\tilde{\alpha}x^2n\}$, we have for all $t\in(t_0, T_n]$,

$$\begin{aligned}\alpha_n'(t)-X'(t) &= -Q(X(t))\{X_n(t)-X(t)\}\\ &\quad +X^0_{\mathrm{A}n}(t-c)\{Q(X_n(t))-Q(X(t))\}\\ &\quad +\{Y_n(t-c)-Y(t-c)\}Q(X(t))\\ &= -M_n(t)\{X_n(t)-X(t)\}+\tilde{R}_n(t),\end{aligned}$$

where $M_n(t)>Q(X(t))>0$ and

$$|\tilde{R}_n(t)|<\{\varepsilon x+2(p+\varepsilon)|X_n(t-c)-X(t-c)|\}Q(X(t)).$$

This follows from (4.31) and because

$$Q(X_n(t))-Q(X(t))=Q'(\theta X(t)+(1-\theta)X_n(t))(X_n(t)-X(t)),$$

for some $\theta\in(0,1)$, and $Q'(x)<0$ for all $x<(X(t)\vee X_n(t))$ and $t\leqslant T_n$. But then

$$\alpha_n'(t)-X'(t)=-M_n(t)\{\alpha_n(t)-X(t)\}+R_n(t),$$

where,

$$\begin{aligned}|R_n(t)| &= |M_n(t)\{\alpha_n(t)-X_n(t)\}+\tilde{R}_n(t)|\\ &< M_n(t)\varepsilon x+\{\varepsilon x+2(p+\varepsilon)\varepsilon x\\ &\qquad +2(p+\varepsilon)|\alpha_n(t-c)-X(t-c)|\}Q(X(t))\\ &\leqslant \{\varepsilon M_n(t)+2\varepsilon Q(X(t))\}x+2(p+\varepsilon)Q(X(t))|\alpha_n(t-c)-X(t-c)|\\ &\equiv U_n(t)x+V_n(t)|\alpha_n(t-c)-X(t-c)|,\end{aligned}$$

for all $t\in(t_0, T_n]$, except on a set of probability at most $\tilde{A}\exp\{-\tilde{\alpha}x^2n\}$. Note that $U_n(t)$ and $V_n(t)$ are positive. Moreover, we have

$$\begin{aligned}U_n(t)+V_n(t)-M_n(t) &= -(1-\varepsilon)M_n(t)+(4\varepsilon+2p)Q(X(t))\\ &< (5\varepsilon+2p-1)Q(X(t))=(-\tfrac{1}{2}+p)Q(X(t))<0.\end{aligned}$$

Case (ii). $p\geqslant\frac{1}{2}$. This implies that

$$X^0_{\mathrm{A}}(\infty)=1-X(\infty)+Y(\infty)>1+X(\infty)(2p-1)\geqslant 1,$$

so that we must have $Q(X(\infty))=0$ and hence $X(\infty)=b_{\mathrm{s}}$. Let $t_0^*>c$ be so large that $Q(X(t_0^*-c))\leqslant 1/(2c)$. Since for $t-c\leqslant s\leqslant t$,

$$X'(s)\leqslant\{1-X(t-c)+2\bar{P}(X(t-c))\}Q(X(t-c)),$$

we have for $t\geqslant t_0^*$,

$$X(t)-X(t-c)\leqslant\tfrac{1}{2}\{1-X(t-c)+2\bar{P}(X(t-c))\}.$$

Hence, for $t\geqslant t_0^*$,

$$\begin{aligned}X^0_{\mathrm{A}}(t-c) &= 1-X(t)+2\bar{P}(X(t-c))\\ &\geqslant \tfrac{1}{2}\{1-X(t-c)+2\bar{P}(X(t-c))\}\\ &\geqslant \tfrac{1}{2}\{1+X(t-c)(2p-1)\}\geqslant\tfrac{1}{2}.\end{aligned}\tag{4.32}$$

By (4.23) we know that $q' = Q'(b_{\mathrm{s}}) < 0$. Take $0 \leqslant \varepsilon \leqslant \frac{1}{2}$ so small that $Q(b_{\mathrm{s}} - \varepsilon k_{\mathrm{h}} y_{\mathrm{h}}) < \frac{1}{54}|q'|$ and $2\varepsilon k_{\mathrm{h}} y_{\mathrm{h}}/d(k_{\mathrm{s}} y_{\mathrm{s}})^2 < \frac{1}{2}|q'|$. Define t_0 as in (4.24) and $\tilde{t}_0 = t_0 \vee t_0^*$. It follows from Lemma 4.3, Theorem 4.1 and Lemma 4.7 that there exist positive numbers $\tilde{A}$, $\tilde{\alpha}$ and $\tilde{C}$ depending on ε, such that for $n = 1, 2, \ldots$ and $0 \leqslant x \leqslant \tilde{C}$,

$$\mathscr{P}\Big(\sup_{t \geqslant 0} |\alpha_n(t) - X_n(t)| \geqslant \tfrac{1}{6}x\Big) \leqslant \tfrac{1}{4}\tilde{A}\exp\{-\tilde{\alpha}x^2 n\}, \tag{4.33a}$$

$$\mathscr{P}\Big(\sup_{t \leqslant \tilde{t}_0} |\alpha_n(t) - X(t)| \geqslant x\Big) \leqslant \tfrac{1}{4}\tilde{A}\exp\{-\tilde{\alpha}x^2 n\}, \tag{4.33b}$$

$$\mathscr{P}\Big(\sup_{t \geqslant \tilde{t}_0} |X(\infty) - X_n(t-c)| \geqslant \varepsilon k_{\mathrm{h}} y_{\mathrm{h}}\Big) \leqslant \tfrac{1}{4}\tilde{A}\exp\{-\tilde{\alpha}x^2 n\}, \tag{4.33c}$$

$$\begin{aligned}&\mathscr{P}\Big(\sup_{t \geqslant \tilde{t}_0} [\,|Y_n(t-c) - Y(t-c)| - 2(p+\varepsilon)|X_n(t-c) - X(t-c)|\,] \geqslant \tfrac{3}{4}x\Big)\\ &\quad \leqslant \tfrac{1}{4}\tilde{A}\exp\{-\tilde{\alpha}x^2 n\}.\end{aligned} \tag{4.33d}$$

Then, except on a set $\tilde{\mathscr{G}}$ of probability not larger than $\tilde{A}\exp\{-\tilde{\alpha}x^2 n\}$, we have for $\tilde{t}_0 < t \leqslant T_n$,

$$\begin{aligned}\alpha_n'(t) - X'(t) &= -Q(X_n(t))\{X_n(t) - X(t)\}\\ &\quad + X_{\mathrm{A}}^0(t-c)\{Q(X_n(t)) - Q(X(t))\}\\ &\quad + \{Y_n(t-c) - Y(t-c)\}Q(X_n(t))\\ &= -M_n(t)\{X_n(t) - X(t)\} + \tilde{R}_n(t),\end{aligned}$$

where $M_n(t) = Q(X_n(t)) - Q'(\theta X_n(t) + (1-\theta)X(t))X_{\mathrm{A}}^0(t-c) \geqslant \frac{1}{4}|q'| > 0$, and

$$\begin{aligned}|\tilde{R}_n(t)| &\leqslant \{\tfrac{3}{4}x + 2(p+\varepsilon)|X_n(t-c) - X(t-c)|\}Q(X_n(t))\\ &\leqslant \tfrac{1}{72}|q'|x + \tfrac{1}{18}|q'|\,|X_n(t-c) - X(t-c)|.\end{aligned}$$

This follows from (4.32), (4.33) and because (cf. (4.23) and (4.33))

$$\begin{aligned}Q'(X_n(t)) &= q' + Q''(\eta X_n(t) + (1-\eta)X(\infty))\{X_n(t) - X(\infty)\}\\ &\leqslant -|q'| + 2\varepsilon k_{\mathrm{h}} y_{\mathrm{h}}/d(k_{\mathrm{s}} y_{\mathrm{s}})^2 \leqslant -\tfrac{1}{2}|q'|.\end{aligned}$$

But then, except on $\tilde{\mathscr{G}}$, we have for $\tilde{t}_0 < t \leqslant T_n$,

$$\alpha_n'(t) - X'(t) = -M_n(t)\{(\alpha_n(t) - X(t)\} + R_n(t),$$

where

$$\begin{aligned}|R_n(t)| &= |M_n(t)\{\alpha_n(t) - X_n(t)\} + \tilde{R}_n(t)|\\ &< \{M_n(t) + \tfrac{1}{12}|q'| + \tfrac{1}{18}|q'|\}\tfrac{1}{6}x + \tfrac{1}{18}|q'|\,|\alpha_n(t-c) + X(t-c)|\\ &\leqslant U_n(t)x + V_n(t)|\alpha_n(t-c) - X(t-c)|\end{aligned}$$

with $U_n(t) = \frac{1}{3}M_n(t)$, $V_n(t) = \frac{1}{3}M_n(t)$. Hence $U_n(t) + V_n(t) - M_n(t) < 0$. This proves assumption (A) and the proof of the theorem is complete. □

Note that, when $p<\frac{1}{2}$, then A and α in Theorem 4.2 depend on p; in case $p\geqslant\frac{1}{2}$, these numbers depend on q'.

Corollary 4.1. *There exist positive numbers A and α such that for $n=1,2,\ldots$ and $x\geqslant 0$,*

$$\mathscr{P}\Big(\sup_{t\geqslant 0}|Y_n(t)-Y(t)|\geqslant x\Big)\leqslant A\exp\{-\alpha x^2 n\}. \tag{4.34}$$

Proof. From (4.20) it can be seen that combination of (4.5), (4.13), the boundedness of P and Theorem 4.2 yields the desired result. □

Having performed the preliminary work, we can now easily prove the uniform convergence of $X_n(t)$ to the solution $X(t)$ of (4.15) for all $t\geqslant 0$, and determine its rate.

Theorem 4.3.

$$\sup_{t\geqslant 0}|X_n(t)-X(t)|=\mathrm{O_P}(n^{-1/2}), \tag{4.35}$$

$$\sup_{t\geqslant 0}|X_n(t)-X(t)|=\mathrm{O}((n^{-1}\log n)^{1/2})\quad a.s. \tag{4.36}$$

Proof. The choice of $x=Bn^{-1/2}$ in Theorem 4.2, where B is a sufficiently large positive constant, immediately proves (4.35).

When, again in Theorem 4.2, $x=B(n^{-1}\log n)^{1/2}$, B positive and large enough, is taken, then application of the Borel-Cantelli lemma yields (4.36). □

In view of Corollary 4.1 we can prove similar results for Y_n and Y:

Corollary 4.2.

$$\sup_{t\geqslant 0}|Y_n(t)-Y(t)|=\mathrm{O_P}(n^{-1/2}), \tag{4.37}$$

$$\sup_{t\geqslant 0}|Y_n(t)-Y(t)|=\mathrm{O}((n^{-1}\log n)^{1/2})\quad a.s. \quad \square \tag{4.38}$$

Finally we can say something about the value of $X(\infty)$. Recall that $X(\infty)=\inf\{z\colon 2\bar{P}(z)-z\leqslant -1\}\wedge b_\mathrm{s}$ (see (4.17)) and that

$$\bar{P}(z)=\int_0^z P(s)\,\mathrm{d}s=(z\wedge b_\mathrm{h})-k_\mathrm{h}y_\mathrm{h}\log\left(\frac{b_\mathrm{h}+k_\mathrm{h}y_\mathrm{h}}{b_\mathrm{h}+k_\mathrm{h}y_\mathrm{h}-(z\wedge b_\mathrm{h})}\right).$$

But $2\bar{P}(z)-z$ is a concave function of z, which equals zero for $z=0$. Hence, if $1+b_\mathrm{h}-2k_\mathrm{h}y_\mathrm{h}\log(1+b_\mathrm{h}(k_\mathrm{h}y_\mathrm{h})^{-1})<0$, then for some $z_0<b_\mathrm{h}$,

$$1+z_0-2k_\mathrm{h}y_\mathrm{h}\log((k_\mathrm{h}y_\mathrm{h}+b_\mathrm{h})/(k_\mathrm{h}y_\mathrm{h}+b_\mathrm{h}-z_0))=0.$$

If not, then

$$\inf\{z\colon 2\bar{P}(z)-z\leqslant -1\}=1+2b_{\mathrm{h}}-2k_{\mathrm{h}}y_{\mathrm{h}}\log(1+b_{\mathrm{h}}(k_{\mathrm{h}}y_{\mathrm{h}})^{-1})\geqslant b_{\mathrm{h}}.$$

Therefore we conclude:

Lemma 4.8. (i) *If* $1+b_{\mathrm{h}}-2k_{\mathrm{h}}y_{\mathrm{h}}\log(1+b_{\mathrm{h}}(k_{\mathrm{h}}y_{\mathrm{h}})^{-1})<0$ *or* $b_{\mathrm{s}}<b_{\mathrm{h}}$, *then*

$$p=P(X(\infty))>0.$$

(ii) *If* $1+b_{\mathrm{h}}-2k_{\mathrm{h}}y_{\mathrm{h}}\log(1+b_{\mathrm{h}}(k_{\mathrm{h}}y_{\mathrm{h}})^{-1})<0$, *then*

$$X(\infty)=\inf\left\{z\colon 1+z-2k_{\mathrm{h}}y_{\mathrm{h}}\log\left(\frac{k_{\mathrm{h}}y_{\mathrm{h}}+b_{\mathrm{h}}}{k_{\mathrm{h}}y_{\mathrm{h}}+b_{\mathrm{h}}-z}\right)=0\right\}\wedge b_{\mathrm{s}}<b_{\mathrm{h}};$$

if $1+b_{\mathrm{h}}-2k_{\mathrm{h}}y_{\mathrm{h}}\log(1+b_{\mathrm{h}}(k_{\mathrm{h}}y_{\mathrm{h}})^{-1})\geqslant 0$, *then*

$$X(\infty)=\{1+2b_{\mathrm{h}}-2k_{\mathrm{h}}y_{\mathrm{h}}\log(1+b_{\mathrm{h}}(k_{\mathrm{h}}y_{\mathrm{h}})^{-1})\}\wedge b_{\mathrm{s}}\geqslant b_{\mathrm{h}}\wedge b_{\mathrm{s}}. \qquad \square$$

5. A central limit theorem

This section is devoted to the derivation of a central limit theorem for the deviation of X_n from X. Because of the results in the foregoing section, we guess that $n^{1/2}$ should be the appropriate scaling factor. We introduce the following notation. For $t\geqslant 0$,

$$V_n(t)=n^{1/2}(X_n(t)-X(t)), \qquad V_n^*(t)=n^{1/2}(\alpha_n(t)-X(t)),$$
$$U_n(t)=n^{1/2}(Y_n(t)-Y(t)),$$

where $\alpha_n=n^{-1}A_{zn}$ as before. In this section convergence in distribution is in $D[0,\infty)$ with respect to the Skorohod metric.

Lemma 5.1.

$$n^{1/2}(X_n(t)-\alpha_n(t))\xrightarrow{\mathscr{D}} W_0(X(t)), \quad \textit{as } n\to\infty, \tag{5.1}$$

where W_0 *is a standard Wiener process.*

Proof. Since N_n^* in (3.13) is a unit Poisson process, we have

$$n^{-1/2}(N_n^*(nu)-nu)\xrightarrow{\mathscr{D}} W_0(u), \tag{5.2}$$

where W_0 is a standard Wiener process. Let $u_n=n^{-1}A_{zn}(t)=\alpha_n(t)$, then

$$\begin{aligned} n^{-1/2}(N_n^*(nu_n)-nu_n) &= n^{-1/2}(N_n(A_{zn}^{-1}(nu_n))-n-nu_n)\\ &= n^{-1/2}(N_n(t)-n-A_{zn}(t))\\ &= n^{1/2}(X_n(t)-\alpha_n(t)). \end{aligned}$$

Moreover, since with probability 1, $u_n\to X(t)$, as $n\to\infty$, the continuity of the paths of W_0 guarantees (5.1). $\square$

Lemma 5.2.

$$
\begin{aligned}
&(V_n(t) - V_n^*(t),\ U_n(t) - 2P(X(t))V_n(t)) \\
&\quad \xrightarrow{\mathscr{D}} \left(W_0(X(t)),\ W_1\left(2\int_c^{X(t)} P(s)(1-P(s))\,\mathrm{d}s\right)\right), \quad \textit{as } n \to \infty, \qquad (5.3)
\end{aligned}
$$

where W_0 *and* W_1 *are independent standard Wiener processes.*

Proof. Since $V_n(t) - V_n^*(t) = n^{1/2}(X_n(t) - \alpha_n(t))$, the weak convergence of the first component was established in Lemma 5.1.

Consider $S_n(u) = n^{-1/2}\sum_{i=1}^{[nu]}(Z_{in} - 2P_{in})$ for $u \in [0, b_{\mathrm{h}}]$ and $n = 1, 2, \ldots$. Analogously to (4.13) we can prove

$$
\max_m \left| \sum_{i=1}^m P_{in}(1-P_{in}) - n\int_0^{m/n} P(s)(1-P(s))\,\mathrm{d}s \right| \leqslant A
$$

for some positive constant A. Hence, it follows from the extension of Donsker's Theorem to sums of independent but not identically distributed random variables (cf. Prohorov, 1956) that

$$
S_n(u) \xrightarrow{\mathscr{D}} W_1\left(2\int_0^u P(s)(1-P(s))\,\mathrm{d}s\right), \quad \text{as } n \to \infty, \qquad (5.4)
$$

where W_1 is a standard Wiener process. In view of (4.13) this means that

$$
n^{-1/2}\left(\sum_{i=1}^{[nu]}(Z_{in} - 2n\bar{P}(u))\right) \xrightarrow{\mathscr{D}} W_1\left(2\int_0^u P(s)(1-P(s))\,\mathrm{d}s\right).
$$

Since $Y_n(t) = n^{-1}\sum_{i=1}^{nX_n(t)} Z_{in}$, $\sup_{t\geqslant 0}|X_n(t) - X(t)| = \mathrm{O}(n^{-1}(\log n)^{1/2}))$ a.s., and $X_n \leqslant b_{\mathrm{s}}$ for all n, the continuity of the sample paths of W_1 ensures that

$$
n^{1/2}(Y_n(t) - 2\bar{P}(X_n(t))) \xrightarrow{\mathscr{D}} W_1\left(2\int_0^{X(t)} P(s)(1-P(s))\,\mathrm{d}s\right).
$$

Now $Y(t) = 2\bar{P}(X(t))$, $\bar{P}'(u) = P(u)$, $U_n(t) = n^{1/2}(Y_n(t) - Y(t))$, $V_n(t) = n^{1/2}(X_n(t) - X(t))$. In view of (4.23) and Theorem 4.2 this implies

$$
U_n(t) - 2P(X(t))V_n(t) \xrightarrow{\mathscr{D}} W_1\left(2\int_c^{X(t)} P(s)(1-P(s))\,\mathrm{d}s\right), \quad \text{as } n \to \infty, \qquad (5.5)
$$

which proves the convergence of the second component in (5.3).

As we pointed out below (3.13), the processes N_n^* and S_n are independent, so that we have joint convergence in (5.2) and (5.4) and hence in (5.3) if W_0 and W_1 are chosen to be independent. □

Lemma 5.3.

$$
\begin{aligned}
\sup_{t\geqslant c}\{&V_n^{*\prime}(t) - [V_n(t)\{-Q(X(t)) + X_{\mathrm{A}}^0(t-c)Q'(X(t))\} \\
&\qquad + U_n(t-c)Q(X(t))]\} \\
&= \mathrm{o}(1) \quad \textit{a.s.} \quad \textit{as } n \to \infty.
\end{aligned}
$$

Proof. In view of (4.23), (4.36) and (4.38), we have as $n\to\infty$ with probability 1, uniformly in $t\geqslant c$,

$$\begin{aligned}
V_n^{*\prime}(t) &= n^{1/2}(-(X_n(t)-X(t))+Y_n(t-c)-Y(t-c))Q(X(t))\\
&\quad +n^{1/2}(1-X_n(t)+Y_n(t-c))(Q(X_n(t))-Q(X(t)))\\
&=-V_n(t)Q(X(t))+U_n(t-c)Q(X(t))\\
&\quad +(1-X_n(t)+Y_n(t-c))Q'(X(t))V_n(t)+\mathrm{o}(1)\\
&=-V_n(t)Q(X(t))+U_n(t-c)Q(X(t))\\
&\quad +(1-X(t)+Y(t-c))Q'(X(t))V_n(t)+\mathrm{o}(1)\\
&=V_n(t)\{-Q(X(t))+X_{\mathrm{A}}^0(t-c)Q'(X(t))\}\\
&\quad +U_n(t-c)Q(X(t))+\mathrm{o}(1). \qquad \square
\end{aligned} \tag{5.6}$$

Lemma 5.4. *The sequence V_n is tight and every weak limit point is in $C[0,\infty)$ a.s.*

Proof. In view of Lemma 5.1, we only have to show that V_n^* is tight and that every weak limit point of V_n^* is in $C[0,\infty)$ a.s. From (5.6) and the boundedness of Q' it follows that

$$\sup_{t\geqslant c}|V_n^{*\prime}(t)|\leqslant A\left\{\sup_{t\geqslant c}|V_n(t)|+\sup_{t\geqslant c}|U_n(t)|\right\}$$

for some positive constant A. Therefore, we see from Theorem 4.2 and Corollary 4.1 that positive $\tilde{A}$ and $\tilde{\alpha}$ exist such that

$$\mathscr{P}\left(\sup_{t\geqslant c}|V_n^{*\prime}(t)|\geqslant x\right)\leqslant \tilde{A}\exp\{-\tilde{\alpha}x^2\}\leqslant\frac{\tilde{A}}{\tilde{\alpha}\mathrm{e}x^2}.$$

Hence, for $\varepsilon>0$, $\eta>0$ and $\delta=(\tilde{\alpha}\mathrm{e}\eta^2\varepsilon/\tilde{A})^{1/2}$,

$$\begin{aligned}
&\mathscr{P}(\sup\{|V_n^*(t)-V_n^*(s)|:\ c\leqslant s<t,\ |t-s|\leqslant\delta\}\geqslant\eta)\\
&\quad\leqslant\mathscr{P}\left(\sup_{t\geqslant c}|V_n^{*\prime}(t)|\geqslant\frac{\eta}{\delta}\right)\leqslant\frac{\tilde{A}\delta^2}{\tilde{\alpha}\mathrm{e}\eta^2}=\varepsilon.
\end{aligned}$$

In view of Pollard (1984, p. 131), this ensures that V_n^* is tight in $D[0,\infty)$ with respect to the Skorohod metric. It also implies, however, that V_n^* is tight in $C[0,\infty)$ with respect to the supremum norm on bounded intervals (cf. Whitt, 1970). The standard subsequence argument shows that every weak limit point of V_n^* is in $C[0,\infty)$ a.s. $\square$

Suppose that V and $\tilde{V}$ are weak limit points of V_n. Assuming our probability space to be sufficiently rich, Lemmas 5.2, 5.3 and a Skorohod construction allow us to redefine V_n, V_n^*, U_n, V, $\tilde{V}$, W_0 and W_1 such that for some subsequences $\{n_k\}$ and $\{\tilde{n}_k\}$, and all t,

$$\lim_{k\to\infty} V_{n_k}(t) = V(t), \qquad \lim_{k\to\infty} V_{\tilde{n}_k}(t) = \tilde{V}(t), \tag{5.7a}$$

$$\lim_{n\to\infty} (V_n(t) - V_n^*(t)) = W_0(X(t)), \tag{5.7b}$$

$$\lim_{n\to\infty} (U_n(t) - 2P(X(t))V_n(t)) = W_1\left(2\int_0^{X(t)} P(s)(1-P(s))\,\mathrm{d}s\right), \tag{5.7c}$$

$$\begin{aligned}\lim_{n\to\infty} \Bigg(V_n^*(t) - \int_c^t [&V_n(s)\{-Q(X(s)) + X_{\mathrm{A}}^0(s-c)Q'(X(s))\} \\ &+ U_n(s-c)Q(X(s))]\,\mathrm{d}s \Bigg) = 0,\end{aligned} \tag{5.7d}$$

with probability 1. Here we have used the continuity of the sample paths of W_0, W_1, V, $\tilde{V}$ (cf. Lemma 5.4) to conclude that almost sure convergence of V_{n_k}, $V_{\tilde{n}_k}$ and $(V_n - V_n^*)$ in the Skorohod metric ensures almost sure convergence in the supremum metric on finite intervals, and hence for all t a.s. Therefore,

$$\begin{aligned} V(t) &= W_0(X(t)) + \lim_{k\to\infty} V_{n_k}^*(t) \\ &= W_0(X(t)) + \int_c^t V(s)\{-Q(X(s)) + X_{\mathrm{A}}^0(s-c)Q'(X(s))\}\,\mathrm{d}s \\ &\quad + \int_c^t \left[W_1\left(\int_0^{X(s-c)} P(u)(1-P(u))\,\mathrm{d}u\right)\right. \\ &\qquad \left. + 2P(X(s-c))V(s-c)\right] Q(X(s))\,\mathrm{d}s \end{aligned}$$

with probability 1, and the same holds with n_k replaced by $\tilde{n}_k$ and V replaced by $\tilde{V}$ throughout. Here the integrals are defined to be zero for $0 \leqslant t < c$.

Consider $F^* : [0,\infty) \times D[0,\infty) \to \mathbb{R}$ defined by

$$\begin{aligned} F^*(t, v) = W_1&\left(2\int_c^{X(t-c)} P(s)(1-P(s))\,\mathrm{d}s\right) Q(X(t)) \\ &+ v(t)\{-Q(X(t) + X_{\mathrm{A}}^0(t-c)Q'(X(t))\} \\ &+ 2v(t-c)P(X(t-c))Q(X(t)), \end{aligned} \tag{5.8}$$

and the system

$$v(t) = \begin{cases} W_0(X(t)) + \displaystyle\int_c^t F^*(s, v)\,\mathrm{d}s, & t \geqslant c, \\ 0, & t < c. \end{cases} \tag{5.9}$$

Then V and $\tilde{V}$ are solutions of (5.9) almost surely. Since for all t, $X(t) \leqslant b_s$, P-almost every realization of $W_0(X(t))$ and $W_1(2\int_0^{X(t)} P(s)(1-P(s))\,\mathrm{d}s)$ are bounded. -

Therefore, because of the boundedness of X, P, $\bar{P}$, Q and Q', and the continuity of the sample paths of W_0 and W_1, the system (5.9) satisfies the conditions needed to have, with probability 1, a unique solution on $[0, \infty)$ (see De Gunst, 1989). Hence $V = \tilde{V}$ a.s. Since V_n is tight, it follows that $V_n \xrightarrow{\mathscr{D}} V$, where V is the unique solution of (5.9) and W_0 and W_1 are independent.

Finally, let $G(t)$ be defined by

$$G(t) = \int_c^t Q(X(s))\, W_1\left(2 \int_c^{X(s-c)} P(u)(1-P(u))\, \mathrm{d}u\right) \mathrm{d}s. \tag{5.10}$$

Obviously, G is a Gaussian process, independent of W_0, with mean zero and covariance

$$\begin{aligned}&\operatorname{Cov}(G(t), G(r))\\ &\quad = 2 \int_c^{t \wedge r} \left\{ [\bar{Q}(s, t) + \bar{Q}(s, r)] Q(X(s)) \int_0^{X(s-c)} P(u)(1-P(u))\, \mathrm{d}u \right\} \mathrm{d}s,\end{aligned}$$

where

$$\bar{Q}(s, t) = \int_s^t Q(X(u))\, \mathrm{d}u.$$

Moreover, let $H(t)$, $A(t)$ and $B(t)$ be defined by

$$H(t) = W_0(X(t)) + G(t), \tag{5.11a}$$

$$A(t) = -Q(X(t)) + X_A^0(t-c) Q'(X(t)), \tag{5.11b}$$

$$B(t) = 2Q(X(t+c)) P(X(t)). \tag{5.11c}$$

Then $V(t)$ can be written as

$$V(t) = H(t) + \int_c^t A(s) V(s)\, \mathrm{d}s + \int_c^{t-c} B(s) V(s)\, \mathrm{d}s. \tag{5.12}$$

Repeated substitution of this expression for V on the right-hand side of (5.12) yields

$$\begin{aligned}V(t) = \sum_{k=0}^{m} \underset{\substack{t_{i+1} < t_i - c\\ i=1,\dots,k-1\\ t_1 < t-c}}{\int \cdots \int} \prod_{i=1}^{k} B(t_i) \exp\left\{ \sum_{i=1}^{k} \int_{t_i + c}^{t_{i-1}} A(u)\, \mathrm{d}u \right\}&\\ \times \left[H(t_k) + \int_c^{t_k} A(s) H(s) \exp\left\{ \int_s^{t_k} A(u)\, \mathrm{d}u \right\} \mathrm{d}s \right] \mathrm{d}t_1 \cdots \mathrm{d}t_k&,\end{aligned} \tag{5.13}$$

where m denotes the integer part of $(t-c)/c$. By defining

$$C(s, t) = \sum_{k=0}^{m} \underset{\substack{t_{i+1} < t_i - c\\ i=1,\dots,k-1\\ t_1 < t-c,\ s < t_k}}{\int \cdots \int} \prod_{i=1}^{k} \left(B(t_i) \exp\left\{ -\int_{t_i}^{t_i + c} A(u)\, \mathrm{d}u \right\} \right) \mathrm{d}t_1 \cdots \mathrm{d}t_k,$$

where $C(s,t)\equiv 1$ for $k=0$, and integrating the term between square brackets in (5.13) by parts, we can modify this expression to get

$$V(t)=\int_c^t C(s,t)\exp\left\{\int_s^t A(u)\,\mathrm{d}u\right\}\mathrm{d}H(s).$$

Hence V is also a Gaussian process with mean zero. It follows from (5.10), (5.11) and the fact that W_0 and W_1 are independent standard Wiener processes that the covariance of V is given by

$$\begin{aligned}\mathrm{Cov}(V(t),V(r))=\int_c^{t\wedge r}\Bigg[&K(s,t)K(s,r)X'(s)\\ &+2[K(s,t)\bar{K}(s,r)+K(s,r)\bar{K}(s,t)]Q(X(s))\\ &\times\int_0^{X(s-c)}P(u)(1-P(u))\,\mathrm{d}u\Bigg]\mathrm{d}s,\end{aligned}\qquad(5.14)$$

where

$$K(s,t)=C(s,t)\exp\left\{\int_s^t A(u)\,\mathrm{d}u\right\},$$

and

$$\bar{K}(s,t)=\int_s^t K(u,t)Q(X(u))\,\mathrm{d}u.$$

Thus we have proved:

Theorem 5.1. *Let W_0 and W_1 be independent standard Wiener processes, let F^* be defined by (5.8), and let V be the unique solution of (5.9). Then V is a Gaussian process with mean zero, covariance given by (5.14) and*

$$V_n\xrightarrow{\mathscr{L}}V,\quad \text{as } n\to\infty,$$

where the convergence is in the sense of the Skorohod metric. □

6. Illustrations

To give the reader an idea what the predicted and actual cell numbers look like, we conclude with some illustrations.

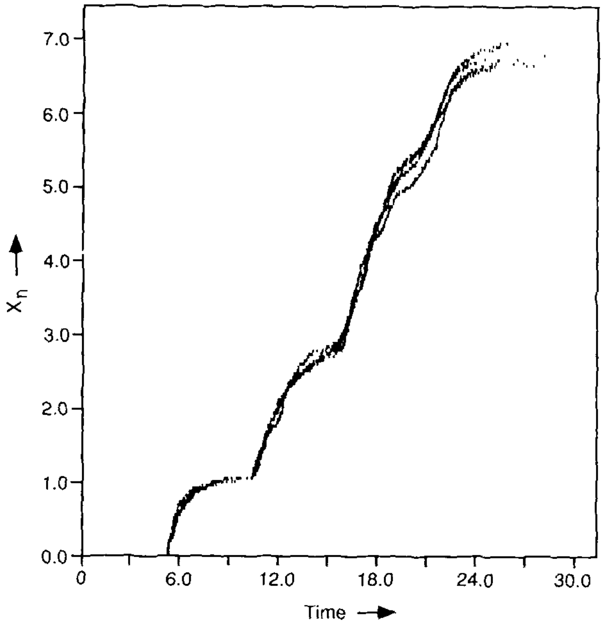

Fig. 1. Four simulations of the stochastic growth process X_n with the same parameter values: $n = 100$, $c = 5$, $d = 1$, $b_s = 10$, $k_s y_s = 0.1$, $b_h = 4$, $k_h y_h = 0.5$.

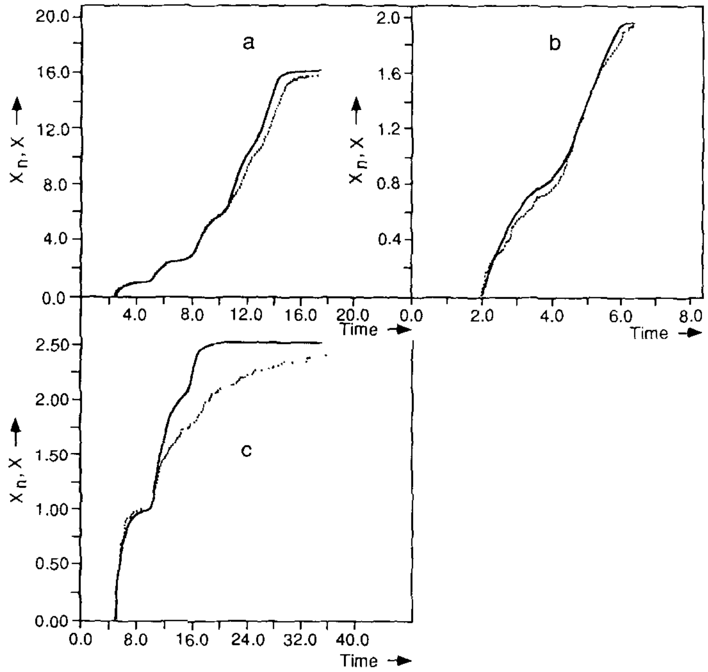

Fig. 2. Three pairs of curves. The dotted curves are computer simulations of X_n; the solid ones are numerically computed simulations of (4.15). Curves of a pair have the same parameter values: (a) $n = 100$, $c = 2.4$, $d = 0.6$, $b_s = 20$, $k_s y_s = 0.01$, $b_h = 10$, $k_h y_h = 1$. Note that $X(\infty) = 1 + 2b_h - 2k_h y_h \log(1 + b_h(k_h y_h)^{-1})$. (b) $n = 100$, $c = 5$, $d = 1$, $b_s = 10$, $k_s y_s = 0.1$, $b_h = 4$, $k_h y_h = 8$. Note that $P(0) < \frac{1}{2}$ and $X(\infty) < b_h$. (c) $b = 100$, $c = 2$, $d = 1$, $b_s = 2$, $k_s y_s = 0.1$, $b_h = 4$, $k_h y_h = 0.1$. Note that $X(\infty) = b_s$.

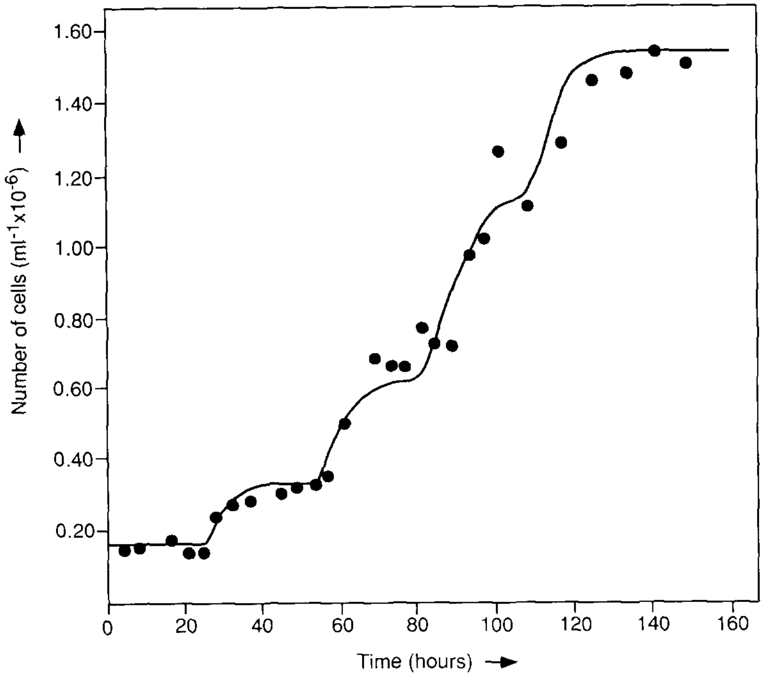

Fig. 3. Experimental data of the growth of a batch culture of tobacco cells (dots). The curve through the data was fitted using numerical solutions of (4.15). The parameter values are $n = 1.625 \times 10^{8}\ \mathrm{l}^{-1}$, $c = 26$ h, $d = 4$ h, $b_s = 29.9$, $k_s = 2.2 \times 10^{-11}$ mol, $y_s = 5.4 \times 10^{10}\ \mathrm{mol}^{-1}$, $b_h = 4.7$, $k_h = 9.7 \times 10^{-17}$ mol, $y_h = 3.4 \times 10^{15}\ \mathrm{mol}^{-1}$.

We have simulated the stochastic growth curve and computed the solution of (4.15) numerically for several sets of parameter values. In Figure 1 four computer simulations of the stochastic process X_n are shown. The parameter values were taken the same for all curves. In Figure 2 three pairs of curves are shown. For the two curves of a pair the same parameter values were used. The dotted curve of a pair was obtained by computer simulation of the stochastic process; the solid one is a numerically computed solution of (4.15). In all cases the initial cell number was 100.

Figures 1 and 2 indicate that under the model with biologically plausible parameters the initial divisions occur synchronously and that this synchrony dies out gradually in time. These features could not be discovered from the available data on population growth of plant cells in batch culture: in order to compare the predictions of the model with experimental data more detailed information was needed. For this reason in the Botanical Laboratory of the University of Leiden a set of experiments was performed in which the increase in cell number of tobacco cells in batch culture was followed in time. The results of one of the experiments are shown in Figure 3. The curve through the data points is a numerical solution of (4.15) where the parameter values were obtained by a least squares method. A statistical test based on a statistical model for the sampling procedure as used in the experiments showed that on the ground of these results there is no reason to reject our model.

For a detailed description of the computational and experimental procedures and of the statistical analysis see De Gunst (1989) and De Gunst et al. (1990).

Acknowledgements

The authors are indebted to K. in 't Hout, E. Pieters and J. Val who did the computations, to K.R. Libbenga, P.A.A. Harkes and E. de Graaff who carried out the experiments and to G.P.G. Hock who drew the figures.

References

M.C.M. de Gunst, A random model for plant cell population growth, CWI Tract No. 58 (Centre for Mathematics and Computer Science, Amsterdam, 1989).

M.C.M. de Gunst, P.A.A. Harkes, J. Val, W.R. van Zwet and K.R. Libbenga, Modelling the growth of a batch culture of plant cells; a corpuscular approach, Enzyme Microb. Technol. 12 (1990) 61–71.

M.C.M. de Gunst and W.R. van Zwet, A non-Markovian model for cell population growth: tail behaviour and duration of the growth process, submitted.

S.N. Ethier and T.G. Kurtz, Markov Processes; Characterization and Convergence (Wiley, New York, 1986).

M. Jacobsen, Statistical analysis of counting processes, Lectures Notes in Statist. No. 12 (Springer, New York, 1982).

P.J. King and H.E. Street, Growth patterns in cell cultures, in: H.E. Street, ed., Plant Tissue and Cell Culture, Botanical Monography 11 (Blackwell, Oxford, 1973).

T.G. Kurtz, Approximation of Population Processes (SIAM, Philadelphia, PA, 1981).

T.G. Kurtz, Gaussian approximations for Markov chains and counting processes, Proc. 44th ISI-session Madrid, 1983, pp. 361–376.

D. Pollard, Convergence of Stochastic Processes (Springer, New York, 1984)

Y.U. Prohorov, Convergence of random processes and limit theorems in probability theory, Theory Probab. Appl. 1 (1956) 157–214.

H.A. Roels, Energetics and Kinetics in Biotechnology (Elsevier, Amsterdam, 1983).

H.E. Street, Cell (suspension) culture techniques, in: H.E. Street, ed., Plant Tissue and Cell Culture, Botanical Monography 11 (Blackwell, Oxford, 1973).

W. Whitt, Weak convergence of probability measures on the function space $C[0, \infty)$, Ann. Math. Statist. 41 (1970) 939–944.