Chapter 13
Fluid Approximations to Markov Decision Processes with Local Transitions

Alexey Piunovskiy and Yi Zhang

13.1 Introduction

Markov decision processes (MDPs) model many practical problems that arise from queueing systems, telecommunication, inventories, and so on, see [6, 7, 15]. The fundamental results about an MDP model are the existence of an optimal policy and the sufficiency of the deterministic stationary policies out of the more general class of randomized history-dependent ones. On the other hand, from practical point of view, it is at least of equal importance to know how to obtain an optimal or nearly optimal policy. It is known that practically, the policy iteration and value iteration procedures fail to cope with MDP models with large state and action spaces. So for random walks, it is often the case that a deterministic continuous model is taken for analysis even when the underlying problem is in stochastic nature. This is called a fluid approximation.

Fluid approximations are widely used to solve practical problems; examples in the contexts of epidemiology and telecommunication can be found in [11, 16], respectively. In inventory control, the well-known (deterministic) economic-order quantity model can be viewed as a fluid approximation, too, cf. [14]. On the one hand, such fluid models can often be solved much more easily than the corresponding stochastic models. On the other hand, in most cases, they are applied without formal analytical justifications, or the justification focuses on the trajectory level, by showing the trajectory of a scaled stochastic model converges in some

A. Piunovskiy (✉) • Y. Zhang
Department of Mathematical Sciences,
University of Liverpool, Liverpool, L69 7ZL, UK
e-mail: piunov@liv.ac.uk; zy1985@liv.ac.uk

D. Hernández-Hernández and A. Minjárez-Sosa (eds.), *Optimization, Control, and Applications of Stochastic Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-0-8176-8337-5_13,

sense to the one of the fluid model, and this is mainly considered for a continuous-time model, see [8–10] and the references therein. For the justification on the objective level, we refer the reader to [3, 4, 13] and the references therein.

In this chapter, we justify (at the level of objective functions) fluid approximations to a discrete-time MDP model with an undiscounted total cost criterion. This is done for an uncontrolled discrete-time model in [12] under more restrictive conditions. The argument is based on [1, 12, 13].

The rest of this chapter is organized as follows. We describe the concerned MDP model in Sect. 13.2 and formulate the main statements in Sect. 13.3, where two sections are devoted to the standard fluid approximation and the refined fluid approximation. We finish this chapter with a conclusion.

13.2 MDP Model

The MDP model under consideration is defined by the following elements:

- $X = \{0, 1, 2, \dots\}$ is the state space.
- A is the action space, which can be an arbitrary non-empty Borel space, whose topology is omitted from the explicit presentation.
- $p(z|x,a)$ is the one-step transition probability, a stochastic kernel on X given $X \times A$ and (Borel) measurable in $a \in A$.
- $r(x,a)$ is the one-step cost, which is (Borel) measurable in $a \in A$.

Assume that the real measurable functions $q^+(y,a), q^-(y,a)$ and $\rho(y,a)$ on $[0,\infty) \times A$ are given such that $q^+(0,a) = q^-(0,a) = \rho(0,a) = 0$, and on $(0,\infty) \times A$, $q^+(y,a) > 0$, $q^-(y,a) > 0$, and $q^+(y,a) + q^-(y,a) \leq 1$. Then we make the MDP model with the absorbing state zero and local transitions only by defining the one-step transition probability and cost via

$$p(z|x,a) = \begin{cases} q^+(x,a), & \text{if } z = x+1; \\ q^-(x,a), & \text{if } z = x-1; \\ 1 - q^+(x,a) - q^-(x,a), & \text{if } z = x; \\ 0 & \text{otherwise,} \end{cases}$$

$$r(x,a) = \rho(x,a).$$

Let $\varphi : X \to A$ be a deterministic stationary policy. For any fixed initial state x and policy φ, the standard canonical construction gives a strategic measure P_x^φ on the space of histories in the form of $x_0, a_0, x_1, a_1, \dots$, and the corresponding expectation is denoted by E_x^φ, see [7]. We denote the controlled process by $\{X_t, t = 0, 1, \dots\}$ and the action process by $\{A_t, t = 0, 1, \dots\}$. Then the MDP model is the following optimization problem, which is well defined after we impose some conditions below:

$$V^\varphi(x) := E_x^\varphi \left[\sum_{t=0}^{\infty} r(X_t, A_t) \right] \to \inf_\varphi,$$

where the infimum is taken over the class of deterministic stationary policies only for simplicity and that under very general conditions, they suffice for the underlying optimization problem, see [1] for more details.

In this chapter, we shall actually scale the above described MDP model such that for any fixed scaling parameter $n = 1,2,\dots$, the elements of the n-MDP model are as follows:

- $X = \{0,1,2,\dots\}$ and A remain as the state and action spaces.
- $$^n p(z|x,a) = \begin{cases} q^+(x/n,a), & \text{if } z = x+1; \\ q^-(x/n,a), & \text{if } z = x-1; \\ 1-q^+(x/n,a)-q^-(x/n,a), & \text{if } z = x; \\ 0 & \text{otherwise} \end{cases}$$
  is the one-step transition probability.
- $^n r(x,a) = \frac{\rho(x/n,a)}{n}$ is the one-step cost, which is measurable in $a \in A$.

The n-MDP model reads

$$^n V^\varphi(x) := E_x^\varphi\left[\sum_{t=0}^{\infty} {}^n r(X_t,A_t)\right] \to \inf_\varphi.$$

Below, we impose some conditions to guarantee that the n-MDP model under consideration is absorbing in the sense of [1]. To be exact, that means given any initial state x, $E_x^\varphi[T_0] < \infty$, where $T_0 := \inf\{t > 0 : X_t = 0\}$. Here and below, the context always makes it clear what the scaling parameter is so that the controlled and action processes in the n-MDP model are still denoted by $\{X_t, t = 0,1,\dots\}$ and $\{A_t, t = 0,1,\dots\}$ for brevity.

The above scaling is called the fluid scaling. Its intuitive meaning together with its importance is now explained via the following example, where an uncontrolled situation is considered for simplicity. Accordingly, simpler denotations are employed. We remark that the example is better understood in the context of telecommunication, where fluid models are widely used to solve satisfactorily practical problems of stochastic nature, see [2, 16] and the reference therein.

Example 13.2.1. Suppose information packets, 1 kilobit (KB) each, arrive at a router (switch) at the (constant) rate $q^+ > 0$ megabit/second (MB/s), and are served at the (constant) rate $q^- > q^+$ MB/s, where $q^+ + q^- \leq 1$. We observe the process up to the moment when the router buffer is empty. Let the holding cost be h per MB per second, so that $\rho(y) = hy$, where y is the amount of information (MB). For simplicity, we consider the uncontrolled model so that the denotation of the policy φ does not appear. One can consider batch arrivals and batch services of 1,000 packets every second; then $n = 1$ and $r(x) = hx = \rho(x)$.

On the other hand, it would be more accurate to consider particular packets; then probabilities q^+ and q^- will be the same, but the time unit is $\frac{1}{1000}$ s, so that the arrival and service rates (MB/s) remain the same. Remembering that, we consider information up to the individual packets (cf. batches) and the time unit is $\frac{1}{1000}$ s, the cost function for the n-model will obviously change: ${}^n r(x) = \frac{hx}{1000}\frac{1}{1000} = \frac{\rho(x/n)}{n}$, where $n = 1{,}000$. □

The goal of this chapter is to estimate (from the above) the differences between

$$ {}^nV^{\varphi}({}^nX_0) := E^{\varphi}_{{}^nX_0}\left[\sum_{t=0}^{\infty} {}^n r(X_t, A_t)\right] $$

and the objective functions of two related deterministic continuous models, namely, the standard fluid model and the refined fluid model, which are simpler to solve. So they are regarded and used as the fluid approximations to the original (stochastic) MDP model, see [11, 16] for examples. In greater detail, under some conditions, we provide explicit upper boundary estimates of the absolute differences between the objective functions of the stochastic and the corresponding fluid models in the initial data, which are understood as the level of accuracy of such fluid approximations. In a nutshell, under the imposed conditions, the absolute differences go to zero as fast as $\frac{1}{n}$, with n being the scaling parameter.

13.3 Main Statements

13.3.1 Standard Fluid Approximation

Firstly, we motivate the standard fluid model by using Example 13.2.1. Then we give its formal definition and obtain its level of accuracy in approximating the n-MDP model.

Example 13.2.1 continued. Consider the situation in Example 13.2.1. The total holding cost of the n-model ${}^nV(x)$ up to the absorption at the state zero satisfies the equation (cf.[12, (10)])

$$ \frac{\rho\left(\frac{x}{n}\right)}{n} + q^+\, {}^nV(x+1) + q^-\, {}^nV(x-1) - (q^+ + q^-)\, {}^nV(x) = 0, $$
$$ x = 1, 2, \ldots; {}^nV(0) = 0. \tag{13.1} $$

Since we measure information in MB, it is reasonable to introduce the function $\hat{v}(y)$ such that ${}^nV(x) = \hat{v}(x/n)$, where $\hat{v}(y)$ is the total holding cost up to the absorption given the initial queue was y MB. After the obvious rearrangements of (13.1), we obtain that $\hat{v}(x/n)$ satisfies

$$ \rho\left(\frac{x}{n}\right) + \frac{q^+\left\{\hat{v}\left(\frac{x}{n}+\frac{1}{n}\right) - \hat{v}\left(\frac{x}{n}\right)\right\}}{\frac{1}{n}} + \frac{q^-\left\{\hat{v}\left(\frac{x}{n}-\frac{1}{n}\right) - \hat{v}\left(\frac{x}{n}\right)\right\}}{\frac{1}{n}} = 0, $$

This is a version of the Euler method for solving the differential equation

$$\rho(y) + (q^+ - q^-)\frac{\mathrm{d}v(y)}{\mathrm{d}y} = 0. \tag{13.2}$$

Thus, we expect that ${}^nV(x) = \hat{v}(x/n) \approx v(x/n)$ at least for a big value of n. □

The above example reveals that as far as the objective function is concerned, the n-MDP model can be approximated by a deterministic continuous model specified by a differential equation, at least for a big value of n. This gives the rise to the following standard fluid model.

The standard fluid model:

$$v^\psi(y_0) := \int_0^\infty \rho(y(\tau), \psi(y(\tau)))\mathrm{d}\tau \to \inf_\psi$$

$$\text{subject to } \frac{\mathrm{d}y}{\mathrm{d}\tau} = q^+(y,\psi(y)) - q^-(y,\psi(y)), \text{ with the given initial state } y(0).$$

Here, ψ is a measurable mapping from $[0,\infty)$ to A. Later, we often omit the argument τ from $y(\tau)$ for brevity. Under the conditions of Theorem 13.3.1 below, it can be seen that

$$v^\psi(y) = \int_0^y \frac{\rho(z,\psi(z))}{q^-(z,\psi(z)) - q^+(z,\psi(z))}\mathrm{d}z, \tag{13.3}$$

cf. (13.2).

Theorem 13.3.1 (cf. Theorem 1 in [12]). *Let $n = 1,2,\dots$ and a policy ψ for the fluid model be fixed, and $\hat{\varphi}$ be given by $\hat{\varphi}(x) := \psi(x/n)$. Suppose*

$$q^-(y,\psi(y)) > \underline{q} > 0, \ \inf_{y>0}\frac{q^-(y,\psi(y))}{q^+(y,\psi(y))} \ge \tilde{\eta} > 1, \ \sup_{y>0}\frac{|\rho(y,\psi(y))|}{\eta^y} < \infty,$$

where $\underline{q} > 0$ is a constant, and $\eta \in (1,\tilde{\eta})$. Then:

(a) $\sup_{x=1,2,\dots}\frac{|{}^nV^{\hat{\varphi}}(x)|}{\eta^{x/n}} < \infty$, *i.e., ${}^nV^{\hat{\varphi}}(\cdot)$ is ${}^n\eta$-bounded, where ${}^n\eta(x) := \eta^{x/n}$.*

(b) If additionally the functions $q^+(y,\psi(y))$, $q^-(y,\psi(y))$, $\rho(y,\psi(y))$ are piecewise continuously differentiable, then for an arbitrarily fixed $\hat{y} \ge 0$

$$\lim_{n\to\infty} \sup_{x\in\{0,1,\dots,[n\hat{y}]\}} |{}^nV^{\hat{\varphi}}(x) - v^\psi(x/n)| = 0,$$

where the function $[\cdot]$ takes the integer part of its argument.

(c) *If furthermore functions* $\rho(y,\psi(y))$, $q^-(y,\psi(y))$, $q^+(y,\psi(y))$ *are continuously differentiable with uniformly bounded derivatives so that* $\sup_{y>0}\left|\frac{d^2 v^\psi(y)}{dy^2}\right| := C < \infty$, *then for any arbitrarily fixed* $\hat{y} \geq 0$,

$$\sup_{x\in\{0,1,\ldots,[n\hat{y}]\}} \left|{}^nV^{\hat{\varphi}}(x) - v^\psi(x/n)\right| \leq \frac{C(3\tilde{\eta}+1)}{2\gamma\tilde{\eta}} \frac{(\hat{y}+1)\tilde{\eta}^{\hat{y}}}{n}.$$

The detailed proof of the above theorem can be found in [12], see the proof of Theorem 1 therein.

The next example shows that the condition $\sup_{y>0} \frac{|\rho(y,\psi(y))|}{\eta^y} < \infty$ in the above theorem is important.

Example 13.3.2 (cf. Example 3 in [12]). Let $\mathbf{A} = [1,2]$, $q^+(y,a) = ad^+$, $q^-(y,a) = ad^-$ for $y > 0$, where $d^- > d^+ > 0$ are fixed numbers such that $2(d^+ + d^-) \leq 1$. Put $\rho(y,a) = a^2\gamma^{y^2}$, where $\gamma > 1$ is a constant. So $\tilde{\eta} = \frac{d^-}{d^+} > 1$.

To solve the fluid model $v^\psi(y) \to \inf_\psi$, we use the dynamic programming approach. One can see that the Bellman function $v^*(y) := \inf_\psi v^\psi(y)$ has the form

$$v^*(y) = \int_0^y \inf_{a\in A} \left\{ \frac{\rho(u,a)}{q^-(u,a) - q^+(u,a)} \right\} du,$$

and satisfies the Bellman equation

$$\inf_{a\in\mathbf{A}} \left\{ \frac{dv^*(y)}{dy} \left[q^+(y,a) - q^-(y,a)\right] + \rho(y,a) \right\} = 0;\ v^*(0) = 0,$$

cf. [13, Lemma 2] and the "incidental" statement in its proof. Here, we remark that the function $\inf_{a\in A} \frac{\rho(u,a)}{q^-(u,a)-q^+(u,a)}$ is universally measurable, see [5, Chap. 7] for more details. Hence, the function

$$v^*(y) = v^{\psi^*}(y) = \int_0^y \frac{\gamma^{u^2}}{d^- - d^+} du$$

is well defined, and $\psi^*(y) \equiv 1$ is the optimal policy.

We notice that the condition $\sup_{y>0} \frac{|\rho(y,\psi^*(y))|}{\eta^y} < \infty$ is not satisfied, while all the other requirements of Theorem 13.3.1 are met.

On the other hand, for the policy given by $\hat{\varphi}(x) = \psi^*(\frac{x}{n}) \equiv 1$ and $n = 1,2,\ldots$, ${}^nV^{\hat{\varphi}}(x) = {}^nE^{\hat{\varphi}}\left[\sum_{t=0}^\infty {}^nr(X_t,A_t)\right]$ satisfies the following equation

$$\frac{\gamma^{(\frac{x}{n})^2}}{n} + d^+\,{}^nV^{\hat{\varphi}}(x+1) + d^-\,{}^nV^{\hat{\varphi}}(x-1) - (d^+ + d^-){}^nV^{\hat{\varphi}}(x) = 0;\ {}^nV^{\hat{\varphi}}(0) = 0,$$

cf. (13.1). But then, this equation does not admit non-negative finitely valued solutions, because if we put ${}^{n}V^{\hat{\varphi}}(0) = 0, {}^{n}V^{\hat{\varphi}}(1) = b$, where $b \in [0,\infty)$ is a non-negative constant, then for any $x = 1,2,\dots,$

$$ {}^{n}V^{\hat{\varphi}}(x) = b\frac{\tilde{\eta}^{x}-1}{\tilde{\eta}-1} - \frac{1}{nd^{+}(\tilde{\eta}-1)}\sum_{j=1}^{x-1}\gamma^{(j/n)^2}(\tilde{\eta}^{x-j}-1), $$

and thus, for a big enough value of x, we obtain ${}^{n}V^{\hat{\varphi}}(x) < 0$. Therefore, Theorem 13.3.1 does not hold; ${}^{n}V^{\hat{\varphi}}(x) = \infty$ for all $x = 1,2,\dots.$ □

13.3.2 *Refined Fluid Approximation*

Under the conditions of Theorem 13.3.1 except for $q^{-}(y,\psi(y)) > \underline{q}$, (13.3) may not hold, which in comparison with (13.2), suggests that the standard fluid approximation may fail to be accurate in this case; since $q^{-}(y)$ can now approach zero, and $q^{+}(y) < q^{-}(y)$, it could happen that the standard fluid model does not get absorbed at the state zero, while for any fixed $n = 1,2,\dots,$ the stochastic process $\{{}^{n}X_{t}, t = 0,1,\dots\}$ indeed gets absorbed at the state zero, so that the standard fluid model and the n-MDP model could behave qualitatively differently. Example 13.3.4 below illustrates this situation. Nevertheless, in this case, the refined fluid model introduced below still approximates well the n-MDP model under the following condition.

Let $\psi(\cdot)$ be a measurable mapping from $[0,\infty)$ to A. We formulate the following condition.

Condition A. (a) $\inf_{y>0}\frac{q^{-}(y,\psi(y))}{q^{+}(y,\psi(y))} \geq \tilde{\eta} > 1$, $\sup_{y>0}\frac{|\rho(y,\psi(y))|}{\{q^{+}(y,\psi(y))+q^{-}(y,\psi(y))\}\eta^{y}} \leq C_1 < \infty$, where $\eta \in (1,\tilde{\eta})$.

(b) For any n, there exists an (n-dependent) constant $K(n)$ such that

$$ {}^{n}l_{W}(x) := K(n)q^{-}(x/n,\psi(x/n))\eta^{x/n} - 1 > 0, x = 1,2,\dots, $$

and

$$ \sup_{x=1,2,\dots}\frac{|\rho(x/n,\psi(x/n))|}{{}^{n}l_{W}(x)} = \sup_{x=1,2,\dots}\frac{|\rho(x/n,\psi(x/n))|}{K(n)q^{-}(x/n,\psi(x/n))\eta^{x/n}-1} < \infty, $$

where $\eta \in (1,\tilde{\eta})$ comes from part (a) of this condition.

(c) There exist points $y_1, y_2, \dots, y_l, \dots$ with $y_l \to \infty$ as $l \to \infty$ such that on each of the intervals $(0,y_1)$, $(y_1,y_2),\dots,$ the function $\frac{\rho(y,\psi(y))}{q^{-}(y,\psi(y))-q^{+}(y,\psi(y))}$ is Lipschitz continuous.

Simple sufficient conditions for Condition A(b) are given below: see Proposition 13.3.1 and its proof.

Refined fluid model:

$$\int_0^\infty \frac{\rho(y,\psi(y))}{q^+(y,\psi(y))+q^-(y,\psi(y))}\mathrm{d}u \to \inf_\psi$$

$$\text{subject to } \frac{\mathrm{d}y}{\mathrm{d}u} = \frac{q^+(y,\psi(y))-q^-(y,\psi(y))}{q^+(y,\psi(y))+q^-(y,\psi(y))}, \text{ with a given initial state } y(0).$$

One can show under Condition A that

$$\tilde{v}^\psi(y_0) := \int_0^\infty \frac{\rho(y,\psi(y))}{q^+(y,\psi(y))+q^-(y,\psi(y))}\mathrm{d}\tau = \int_0^{y_0} \frac{\rho(z,\psi(z))}{q^-(z,\psi(z))-q^+(z,\psi(z))}\mathrm{d}z.$$

Note that $\frac{q^+(y,\psi(y))-q^-(y,\psi(y))}{q^+(y,\psi(y))+q^-(y,\psi(y))} \leq \frac{1-\tilde{\eta}}{1+\tilde{\eta}} < 0$, so that $y(\cdot)$ in the fluid model is absorbed at the state zero in finite time.

Theorem 13.3.2. *Let $n = 1,2,\ldots$ and $\hat{y} > 0$ be fixed, ψ a measurable mapping from $[0,\infty)$ to A, and $\hat{\varphi}(\cdot)$ given by $\hat{\varphi}(x) := \psi(x/n)$. Suppose Condition A is satisfied for ψ so that there exist an integer L such that the function $\frac{\rho(y,\psi(y))}{q^-(y,\psi(y))-q^+(y,\psi(y))}$ is Lipschitz continuous with the common Lipschitz constant D on the intervals $(0,y_1)$, (y_1,y_2), $\ldots$, (y_{L-1},y_L), (y_l,y_{L+1}) with $y_L < \hat{y}+1 \leq y_{L+1}$. Then*

$$\sup_{x\in\{0,1,\ldots,[\hat{y}n]\}} \left| {}^nV^{\hat{\varphi}}(x) - \tilde{v}^\psi(x/n)\right| \leq \frac{\varepsilon(n)}{2}.$$

Here and below, we put

$$\begin{aligned}
\varepsilon(n) &:= \frac{2K_1}{n} + \frac{2K_2}{\tilde{\eta}^n} + 2K_3(\eta^{1/n}-1),\\
K_1 &:= \frac{\tilde{\eta}+1}{\tilde{\eta}-1}[D(\hat{y}+1)+3C_1L\eta^{\hat{y}+1}],\\
K_2 &:= \frac{\tilde{\eta}+1}{\tilde{\eta}-1}C_1\left[1+\frac{2(\tilde{\eta}+1)}{(\tilde{\eta}-1)\ln\eta}\right]\frac{\eta^{\hat{y}+1}\tilde{\eta}^2}{\tilde{\eta}-\eta},\\
K_3 &:= \left(\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)^2\frac{3C_1L\eta^{\hat{y}+1}}{\ln\eta}.
\end{aligned}$$

Proof. Consider an n-birth-and-death process model whose birth rate is $nq^+(x/n,\hat{\varphi}(x))$, death rate is $nq^-(x/n,\hat{\varphi}(x))$ and cost rate is $\rho(x/n,\hat{\varphi}(x))$. The underlying process is denoted by $\{Y_t, t\geq 0\}$. Then we have

$${}^nW^{\hat{\varphi}}(x) := E_x^{\hat{\varphi}}\left[\int_0^\infty \rho(Y_t/n,\hat{\varphi}(Y_t))\mathrm{d}t\right] = {}^nV^{\hat{\varphi}}(x),\ x=0,1,\ldots.$$

Indeed, both $V^{\hat{\varphi}}(x)$ and ${}^nW^{\hat{\varphi}}(x)$ are given by the unique ${}^n\eta$-bounded solution to the equation

$$\begin{aligned} 0 &= \rho(x/n,\hat{\varphi}(x)) + V(x+1)nq^+(x/n,\hat{\varphi}(x)) + nq^-(x/n,\hat{\varphi}(x))V(x-1) \\ &\quad -n(q^+(x/n,\hat{\varphi}(x)) + q^-(x/n,\hat{\varphi}(x)))V(x), x = 1,2,\ldots; \\ 0 &= V(0) \end{aligned}$$

by Remark 13.3.1, [12, Lem. 1, Lem. 2, Lem. 4] and [13, Lem. 1]. See also [1, Chap. 7].

It remains to apply [13, Thm.1]. □

For any $n = 1,2,\ldots$, let us consider a subclass of deterministic stationary policies ${}^n\Pi$ whose elements are φ such that the following are satisfied:

$$\inf_{x=1,2,\ldots} \frac{q^-(x/n,\varphi(x))}{q^+(x/n,\varphi(x))} \geq \tilde{\eta} > 1,$$

$$\sup_{x=1,2,\ldots} \frac{|\rho(x/n,\varphi(x))|}{\{q^+(x/n,\varphi(x)) + q^-(x/n,\varphi(x))\}\eta^x} \leq C_1 < \infty,$$

where $\eta \in (1,\tilde{\eta})$, and there exists an ($\varphi,n$-dependent) constant $K^{\varphi}(n)$ satisfying

$$K^{\varphi}(n)q^-(x/n,\varphi(x))\eta^{x/n} - 1 > 0,\ x = 1,2,\ldots,$$

and

$$\sup_{x=1,2,\ldots} \frac{|\rho(x/n,\varphi(x))|}{K^{\varphi}(n)q^-(x/n,\varphi(x))\eta^{x/n} - 1} < \infty.$$

Note that if there exists a ψ satisfying Condition A, then the set ${}^n\Pi$ is non-empty as $\hat{\varphi}(x) := \psi(x/n) \in {}^n\Pi$ for all $n = 1,2,\ldots$. Under Condition B below, for any $n = 1,2,\ldots$, ${}^n\Pi$ coincides with the whole class of deterministic stationary policies, see Proposition 13.3.1 below.

Remark 13.3.1. For any fixed $n = 1,2,\ldots$, one can verify that under a fixed policy $\varphi \in {}^n\Pi$, the n-MDP model admits a Lyapunov function

$${}^nl_L(x) := D\eta^{x/n}, x = 1,2,\ldots, {}^nl_L(0) := 2,$$

where $D \geq 1$ is a big enough constant, and a weight function

$${}^nl_W(x) := K^{\varphi}(n)q^-(x/n,\varphi(x))\eta^{x/n} - 1, x = 1,2,\ldots, {}^nl_2(0) := 1,$$

cf. [1, Chap.7] and [12, Con.1].

Condition B. (a) $\inf_{y>0,a\in A}\frac{q^-(y,a)}{q^+(y,a)} \geq \tilde{\eta} > 1$, $\sup_{y>0,a\in A}\frac{|\rho(y,a)|}{(q^+(y,a)+q^-(y,a))\eta^y} \leq C_1 < \infty$, where $\eta \in (1,\tilde{\eta})$.
(b) $\liminf_{y\to\infty}\inf_{a\in A} q^-(y,a) > 0$.

Proposition 13.3.1. *Suppose Condition B is satisfied. Then for any deterministic stationary policy φ, it holds that $\varphi \in {}^n\Pi$ for all $n = 1,2,\ldots$.*

Proof. Let $n = 1,2,\ldots$ be fixed and consider an arbitrarily fixed φ. Then under Condition B, there exists a $\zeta > 0$ such that $q^-(x/n,\varphi(x)) > \zeta$, $x = 1,2,\ldots$. So there exists a constant $K^\varphi(n) > 0$ such that $q^-(x/n,\varphi(x))\eta^{x/n} - \frac{1}{K^\varphi(n)} > \tilde{\zeta}, x = 1,2,\ldots$ where $\tilde{\zeta} > 0$. Thus,

$$K^\varphi(n)q^-(x/n,\varphi(x))\eta^{x/n} - 1 > 0,$$

and

$$\begin{aligned}
\sup_{x=1,2,\ldots}\frac{|\rho(\frac{x}{n},\varphi(x))|}{K^\varphi(n)q^-(\frac{x}{n},\varphi(x))\eta^{\frac{x}{n}}-1} &= \frac{1}{K^\varphi(n)}\sup_{x=1,2,\ldots}\frac{|\rho(\frac{x}{n},\varphi(x))|}{q^-(\frac{x}{n},\varphi(x))\eta^{\frac{x}{n}}-\frac{1}{K^\varphi(n)}}\\
&\leq \frac{1}{K^\varphi(n)}\sup_{x=1,2,\ldots}\frac{|\rho(x/n,\varphi(x))|}{(q^+(x/n,\varphi(x))+q^-(x/n,\varphi(x)))\eta^{x/n}}\\
&\quad\times\sup_{x=1,2,\ldots}\frac{2q^-(x/n,\varphi(x))\eta^{x/n}}{q^-(x/n,\varphi(x))\eta^{x/n}-\frac{1}{K^\varphi(n)}}\\
&\leq \frac{2C_1}{K^\varphi(n)}\sup_{x=1,2,\ldots}\frac{q^-(x/n,\varphi(x))\eta^{x/n}}{q^-(x/n,\varphi(x))\eta^{x/n}-\frac{1}{K^\varphi(n)}} < \infty.
\end{aligned}$$

The other requirements for a policy to be in ${}^m\Pi$ are satisfied by φ is evident. □

Theorem 13.3.3. *Suppose the policy ψ^* solves the refined fluid model and satisfies Condition A. Then for any fixed $n = 1,2,\ldots,$*

$$\sup_{x\in\{0,1,\ldots,[n\hat{y}]\}}\left|{}^nV^{\varphi^*}(x) - \inf_{\varphi\in{}^n\Pi}{}^nV^\varphi(x)\right| \leq \varepsilon(n),$$

where $\varphi^*(x) := \psi^*(x/n), x = 0,1,2,\ldots$.

Proof. It can be shown that $\forall\ \varphi \in {}^n\Pi$, ${}^nV^\varphi(x) \geq \tilde{v}^{\psi^*}(x/n) - \frac{\varepsilon(n)}{2}$. The proof is similar to the one of Theorem 13.3.2. On the other hand, from Theorem 13.3.2, we have $\inf_{\varphi\in{}^n\Pi}{}^nV^\varphi(x) \leq {}^nV^{\varphi^*} \leq \tilde{v}^{\psi^*}(x/n) + \frac{\varepsilon(n)}{2} \leq \inf_{\varphi\in{}^n\Pi}{}^nV^\varphi(x) + \varepsilon(n)$. □

The above theorem asserts that if one solves the refined fluid model and obtains the optimal policy ψ^*, then the policy given by $\varphi^*(x) = \psi^*(x/n)$ is $\varepsilon(n)$-optimal in the underlying n-MDP, and $\varepsilon(n)$ goes to zero as n grows large.

The next proposition comes from [13, Lem.3].

Proposition 13.3.2. *Under Condition B, assume that there exist finite intervals* $(0,y'_1)$, $(y'_1,y'_2),\ldots$, *with* $\lim_{j\to\infty} y'_j=\infty$, *such that on each of them, the function* $\frac{\rho(y,a)}{q^-(y,a)-q^+(y,a)}$ *is Lipschitz continuous with respect to* y *for each* $a\in A$, *and the Lipschitz constants are* a*-independent. Then for any fixed* $\hat{y}>0$, *there exists an* ψ^* *satisfying Condition A and solving the refined fluid model on* $[0,\hat{y}]$.

Now, we give an example where the main results (Theorems 13.3.2 and 13.3.3) of this work are applicable. In fact, by Propositions 13.3.1 and 13.3.2, it suffices to verify Condition B and the other condition of Proposition 13.3.2.

Example 13.3.3. Consider a discrete-time single-server queueing system, where during each time step, the probability of having an arrival is given by the function $q^+(y,a)=\frac{a}{2y+2+a}$, and the probability of having a service completion (given there is at least one job) is given by the function $q^-(y,a)=\frac{2y+2}{2y+2+a}$, where $y>0$ and $a\in A:=[\frac{1}{2},1]$. Suppose the cost function is given by $\rho(y,a)=2y-a$, which means that we aim at minimizing the holding cost, which is incurred at a rate of 2£ per time step, while admitting more jobs is encouraged. The state zero is taken as the absorbing state, so that $q^-(0,0)=q^+(0,0)=\rho(0,0)=0$.

It is easy to see that

$$\inf_{y>0,a\in[\frac{1}{2},1]}\frac{q^-(y,a)}{q^+(y,a)}=\inf_{y>0,a\in[\frac{1}{2},1]}\frac{2+2y}{a}=2=:\tilde{\eta}>1,$$

$$\sup_{y>0,a\in[\frac{1}{2},1]}\frac{|\rho(y,a)|}{(q^+(y,a)+q^-(y,a))(1.5)^y}=\frac{|\rho(y,a)|}{(1.5)^y}<\infty,$$

$$1<\eta:=1.5<\tilde{\eta},$$

$$\liminf_{y\to\infty}\inf_{a\in[\frac{1}{2},1]}q^-(y,a)>0,$$

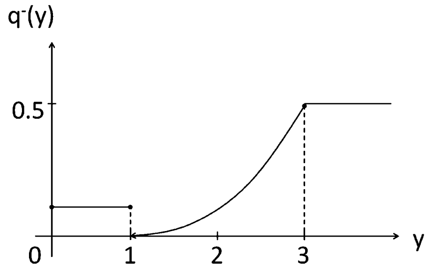
and the function given by

$$\frac{\rho(y,a)}{q^-(y,a)-q^+(y,a)}=\frac{(2y-a)(2y+2+a)}{2y+2-a}=2y+a-\frac{4a}{2y+2-a}$$

is obviously Lipschitz in $y>0$ for any $a\in[\frac{1}{2},1]$. Hence, Condition B and the other condition of Proposition 13.3.2 are satisfied by this example. □

The next example indicates that the condition $q^-(y,\psi(y))>\underline{q}>0$ is important for the standard fluid model to approximate the underlying MDP model accurately. It also illustrates that the standard and the refined fluid models can behave qualitatively differently.

Fig. 13.1 The graph of $q^-(y)$



Example 13.3.4 (cf. Example 1 in [13]). For brevity, we deal with an uncontrolled model, i.e., A is taken as a singleton, so that denotations such as $q^-(y), q^+(y)\rho(y)$ are used for brevity. We put

$$q^-(y) = 0.1I\{y \in (0,1]\} + 0.125(y-1)^2 I\{y \in (1,3]\} + 0.5I\{y > 3\},$$
$$q^+(y) = 0.2q^-(y),\ \rho(y) = 8q^-(y).$$

Clearly, $q^-(y)$ is not separated from zero, see Fig. 13.1, while Condition A is satisfied.

For the original fluid model, we have $\frac{\mathrm{d}y}{\mathrm{d}\tau} = -0.1\,(y-1)^2$, and, if the initial state $y_0 = 2$, then $y(\tau) = 1 + \frac{10}{\tau+10}$, so that $\lim_{\tau\to\infty} y(\tau) = 1$.

On the other hand, since $q^-(y), q^+(y) > 0$ for $y > 0$, and there is a negative trend, the state process X_t in the n-stochastic model starting from ${}^nX_0/n = y_0 = 2$ will be absorbed at zero, see [1, Lem. 7.2, Def. 7.4], while the moment of the absorbtion is postponed for later and later as $n \to \infty$ because the process spends more and more time in the neighborhood of 1, see Figs. 13.2 and 13.3.

When using the original fluid model, we have

$$v(2) = \int_0^\infty \rho(y(\tau))\mathrm{d}\tau = 10 = \lim_{T\to\infty}\lim_{n\to\infty} E_{2n}\left[\sum_{t=1}^\infty I\{t/n \le T\}\, {}^nr(X_{t-1}, A_t)\right].$$

When using the refined fluid model, we can calculate $\frac{\rho(y)}{q^+(y)+q^-(y)} = \frac{8}{1.2}$ for $y > 0$ and $y(u) = 2 - \frac{2}{3}u$, so that the process in the refined fluid model is absorbed at the state zero at the time moment $u = 3$. Therefore,

$$\lim_{n\to\infty} {}^nV(2n) = \tilde{v}(2) = \int_0^\infty \hat{\rho}(y(u))\mathrm{d}u = \int_0^3 \frac{8}{1.2}\mathrm{d}u = 20$$
$$= \lim_{n\to\infty}\lim_{T\to\infty} E_{2n}\left[\sum_{t=1}^\infty I\{t/n \le T\}\, {}^nr(X_{t-1}, A_t)\right] \neq v(2).$$

So the standard fluid model fails to be accurate in this example. □

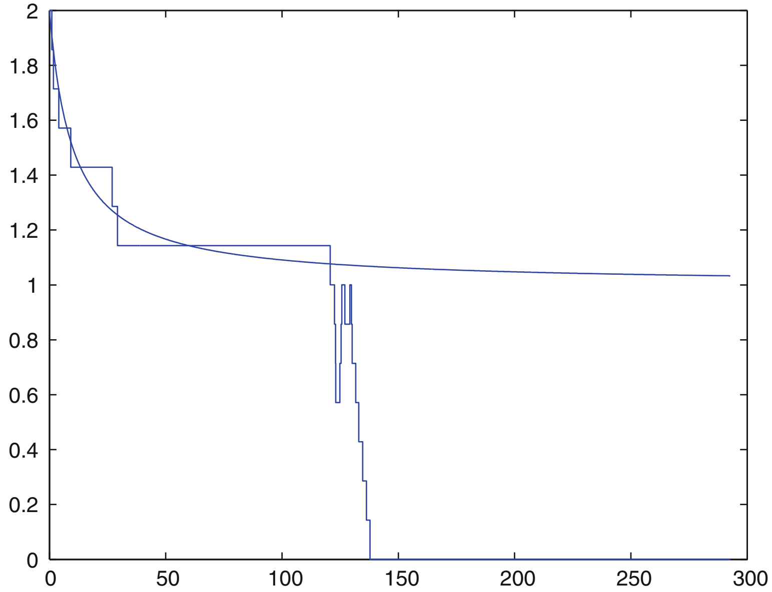


Fig. 13.2 The state process in the n-stochastic model and its fluid approximation, $n = 7$

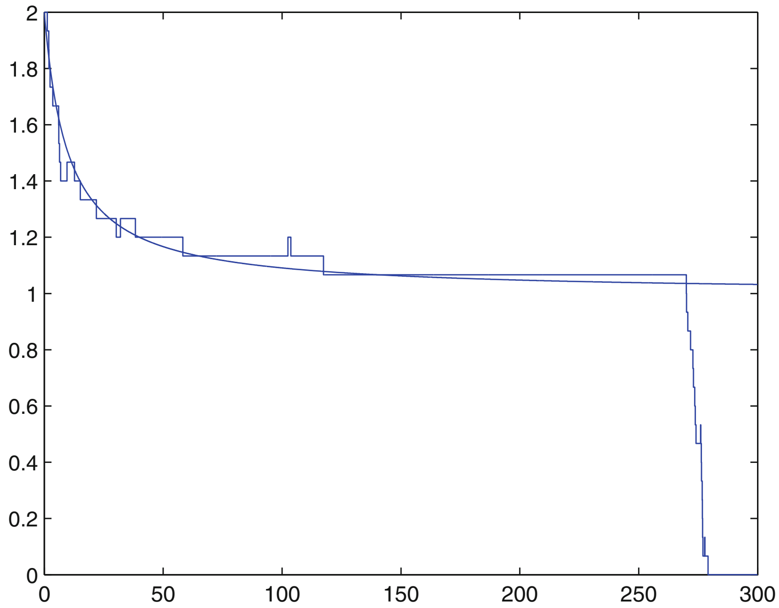


Fig. 13.3 The state process in the n-stochastic model and its fluid approximation, $n = 15$

13.4 Conclusion

In this chapter, the convergence of the objective function of a scaled absorbing MDP model, with a total undiscounted cost, to the one of the (standard and refined) fluid model is shown. The upper boundary estimate of the rate of convergence is presented based on the initial data, which is of order $\frac{1}{n}$, where n is the scaling parameter. Hence, the level of accuracy of the fluid approximation is obtained. By examples, we also show that the standard fluid model may fail to approximate the n-MDP model, while the refined fluid model is still accurate.

Acknowledgements Mr. Mantas Vykertas kindly helped us improve the English presentation of this chapter. We thank the referee for valuable comments, too.

References

1. Altman, E.: Constrained Markov Decision Processes. Chapman and Hall/CRC, Boca Raton (1999)
2. Avrachenkov, K., Ayesta, U., Piunovskiy, A.: Convergence of trajectories and optimal buffer sizing for AIMD congestion control. Perform. Evaluation. **67**, 501–527 (2010)
3. Avrachenkov, K., Piunovskiy, A., Zhang, Y.: Asymptotic fluid optimality and efficiency of tracking policy for bandwidth-sharing networks. J. Appl. Probab. **48**, 90–113 (2011)
4. Bäuerle, N.: Optimal control of queueing networks: an approach via fluid models. Adv. Appl. Prob. **34**, 313–328 (2002)
5. Bertsekas, D., Shreve, S.: Stochastic Optimal Control. Academic Press, NY (1978)
6. Jacko, P., Sansó, B.: Optimal anticipative congestion control of flows with time-varying input stream. Perform. Evaluation. **69**, 86–101 (2012)
7. Hernández-Lerma, O., Lasserre, J.: Discrete-time Markov Control Processes. Springer-Verlag, NY (1996)
8. Dai, J.: On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. Ann. Appl. Prob. **5**, 49–77 (1995)
9. Foss, S., Kovalevskii, A.: A stability criterion via fluid limits and its application to a Polling system. Queueing. Syst. **32**, 131–168 (1999)
10. Mandelbaum, A., Pats, G.: State-dependent queues: approximations and applications. In Kelly, F., Williams, R. (eds.) Stochastic Networks, pp. 239–282. Springer, NY (1995)
11. Piunovskiy, A., Clancy, D.: An explicit optimal intervention policy for a deterministic epidemic model. Optim. Contr. Appl. Met. **29**, 413–428 (2008)
12. Piunovskiy, A.: Random walk, birth-and-death process and their fluid approximations: absorbing case. Math. Meth. Oper. Res. **70**, 285–312 (2009)
13. Piunovskiy, A., Zhang, Y.: Accuracy of fluid approximations to controlled Birth-and-Death processes: absorbing case. Math. Meth. Oper. Res. **73**, 159–187 (2011)
14. Piunovskiy, A., Zhang, Y.: On the fluid approximations of a class of general inventory level-dependent EOQ and EPQ models. Adv. Oper. Res. (2011) doi: 10.1155/2011/301205
15. Puterman, M.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, NY (1994)
16. Zhang, Y., Piunovskiy, A., Ayesta, U., Avrachenkov, K.: Convergence of trajectories and optimal buffer sizing for MIMD congestion control. Com. Com. **33**, 149–159 (2010)