On the Dynamics of Evolutionary Games

María Teresa Gallegos
Fakultät fűr Mathematik und Informatik
Universität Passau
D-94030 Passau
Germany

Abstract: We establish new and extend some existing results on stationary distributions and first passage times of the Markov chain describing an evolutionary game. We derive these results from new lower and upper bounds for the quotient of the equilibrium probabilities of two events of an irreducible Markov chain. These bounds are in terms of the transition matrix and easy to compute.

Key Words: Evolutionary games, Markov chains, passage times, stationary distributions, invariant measures, stochastic bounds

AMS 1991 Subject Classification: Primary 60J10; secondary 90D12, 93E05

1 Introduction

Loosely speaking, a branch of evolutionary game theory tries to explain why in games with multiple Nash equilibria some of them are preferred to others. A typical example of such a game is a *coordination game* which arises when two players have to collaborate in order to achieve a commonly desired outcome, but in which neither player will benefit from his efforts if his partner does not do his part. The idea behind evolutionary theory is that a combination of random experimentation or mutation together with the myopic attempts of players to coordinate with those around them creates *evolutionary* forces which lead the players to coordinate on the risk dominant equilibrium (see Harsanyi and Selten [3]). There is a rapidly growing literature on the subject and we refer the reader to the references cited in Canals and Vega–Redondo [1].

Kandori, Mailath and Rob [5] provide an elegant characterisation of the long run influence of evolutionary forces. The basis of the game theoretical part of the present paper, see Sections 4, 5, and 6, is the model established in this reference. Using the theoretical results developed in the theoretical Markov chain part of the present communication, Sections 2 and 3, we give an analysis of the long run behaviour of the corresponding Markov chain. Our main results are the speeds of convergence of both, the entries of the stationary distribution, and the first passage time to interesting subsets, cf. Theorems 5.4 and Theorem 6.2.

Our contribution to the theory of Markov chains may be summarize as follows: Let Q be an irreducible transition matrix on a finite state space E and let π be its stationary distribution. For two (disjoint) events $A, B \subset E$, we first present upper and lower bounds for the quotient $\frac{\pi(A)}{\pi(B)}$. The estimates are expressed in terms of (partial) row sums of the transition matrix. Although the given bounds may be useful for numerical applications, we focus on Markov kernels with *parametric entries*. Moreover, we do not make any structural assumption on the Markov kernel Q. With these building blocks at hand, we give a complete analysis of the long run behaviour of a Markov chain emerging in an important subset of the 2×2 symmetric *evolutionary games*. This subset includes both the coordination games and the games with a dominant strategy.

1.1 General notations

Let $\mathbf{N} = \{0, 1, 2, \cdots\}$ be the set of natural numbers and let E be a finite set.

(a) Let $m \in \mathbf{N}$. Let α_1 and α_2 be two distributions on $0..m$. The expression $\alpha_1 \preceq \alpha_2$ denotes that α_1 is stochastically smaller than the distribution α_2.

(b) For a matrix $Q \in \mathbf{R}^{E \times E}$ and a subset $D \subseteq E$, we denote by $Q_{|D \times D}$ the restriction of the matrix Q to the set $D \times D$.

(c) Let α be a distribution on E and let D be a subset of E such that $\alpha(D) > 0$. We denote by α_D the probability vector $(\frac{\alpha(d)}{\alpha(D)})_{d \in D}$. The context defines if α_D is a probability vector defined on E with support D or if it is defined on D. Moreover, S_α denotes the support of the distribution α.

(d) For a subset D of E, we denote by $\complement D$ the complement of the set D.

(e) The symbol $\mathbf{1}$ denotes a column vector with all entries equal to the scalar 1. Let $D \subseteq E$. The symbol $\mathbf{1}_D$ denotes the column vector with entries 1 on D and 0 elsewhere. The dimension is defined by the context.

(f) The symbol δ_x denotes the Dirac distribution at an element $x \in E$.

(g) For a transition matrix $K \in \mathbf{R}^{E \times E}$, we denote by K^n the n–step transition matrix.

2 Estimates for the quotient of the equilibrium probabilities of two events

In this section, we develop bounds of general interest for the quotient of the stationary distributions of any two events contained in E. To our knowledge, these results are new, apart from the particular case of Corollary 2.3(b) which appears in reference [2].

Before giving our first proposition recall that, for two nonempty subsets B, C of E, the family $(Q(c, B))_{c \in C}$ induces a preorder on C given by $c_1 \leq c_2$, if $Q(c_1, B) \leq Q(c_2, B)$.

2.1 Proposition

Let B and C be two nonempty, disjoint subsets of E and let $\leq_B$ be the order on B induced by the family $(Q(b, B))_{b \in B}$ and $\leq_C$ be the order on C induced by the family $(Q(c, B))_{c \in B}$. Moreover, let α_1 be a distribution on the (ordered) set $(B, \leq_B)$ such that $\alpha_1 \preceq (\pi_B(b))_{b \in B}$ and let α_2 be a distribution on the (ordered) set $(C, \leq_C)$ such that $\alpha_2 \preceq (\pi_C(c))_{c \in C}$. We have

$$1 - \min_{b \in B} \ Q(b, B) > 0. \tag{1}$$

Hence, the quotients

$$\frac{\sum_{c \in C} \pi_C(c) \ Q(c, B)}{1 - \sum_{b \in B} \pi_B(b) \ Q(b, B)}, \quad \frac{\sum_{c \in C} \alpha_2(c) \ Q(c, B)}{1 - \sum_{b \in B} \alpha_1(b) \ Q(b, B)}, \text{ and } \frac{\min_{c \in C} \ Q(c, B)}{1 - \min_{b \in B} \ Q(b, B)} \tag{2}$$

are well defined and we have

$$\begin{aligned} \frac{\pi(B)}{\pi(C)} &\geq \frac{\sum_{c \in C} \pi_C(c) \ Q(c, B)}{1 - \sum_{b \in B} \pi_B(b) \ Q(b, B)} && (3) \\ &\geq \frac{\sum_{c \in C} \alpha_2(c) \ Q(c, B)}{1 - \sum_{b \in B} \alpha_1(b) \ Q(b, B)} && (4) \\ &\geq \frac{\min_{c \in C} \ \ Q(c, B)}{1 - \min_{b \in B} \ Q(b, B)} && (5) \\ &=: L(B, C). \end{aligned}$$

If $B \cup C = E$ then equality holds in (3).

By interchanging the roles of B and C, we obtain upper bounds for $\frac{\pi(B)}{\pi(C)}$, i.e.,

$$\begin{aligned}
\frac{\pi(B)}{\pi(C)} &\leq \frac{1-\sum_{c\in C}\pi_C(c)\ Q(c,C)}{\sum_{b\in B}\pi_B(b)\ Q(b,C)} && (6)\\
&\leq \frac{1-\sum_{c\in C}\alpha_2(c)\ Q(c,C)}{\sum_{b\in B}\alpha_1(b)\ Q(b,B)} \\
&\leq \frac{1-\min_{c\in C}\ \ Q(c,C)}{\min_{b\in B}\ Q(b,C)} && (7)\\
&= \frac{1}{L(C,B)} \\
&=: U(B,C).
\end{aligned}$$

Proof. Claim (1) is a consequence of the irreducibility of Q and it implies plainly that the last quotient in (2) is well defined. It is well known that, for two distributions β_1 and β_2 on an ordered space $(D,\leq)$ we have $\beta_1 \preceq \beta_2$ if and only if, for all increasing functions $f: D \to D$, we have $\sum_{d\in D}\beta_1(d)f(d) \leq \sum_{d\in D}\beta_2(d)f(d)$. Using this fact together with the hypotheses $\delta_{\operatorname{argmin} Q(b,B)} \preceq \alpha_1 \preceq (\pi_B(b))_{b\in B}$ and $\delta_{\operatorname{argmin} Q(c,B)} \preceq \alpha_c \preceq (\pi_C(c))_{c\in C}$ the first two quotients in (2) are well defined as a direct consequence of the definition of the last one. The same argument proves (4) and (5) from (3). For the proof of (3), note that the equality $(\pi Q)(B) = \pi(B)$ may be written as

$$\pi(B) = \sum_{b\in B}\pi(b)\ Q(b,B) + \sum_{c\in C}\pi(c)\ Q(c,B) + \sum_{d\in \complement(B\cup C)}\pi(d)\ Q(d,B).$$

Hence, we have

$$\sum_{b\in B}\pi(b)\ (1-Q(b,B)) \geq \sum_{c\in C}\pi(c)\ Q(c,B),$$

i.e.,

$$\pi(B)\sum_{b\in B}\pi_B(b)\ (1-Q(b,B)) \geq \pi(C)\sum_{c\in C}\pi_C(c)\ Q(c,B).$$

The claim (3) follows now from the inequality above by a simple algebraic operation. □

2.2 Remarks

(a) Clearly, the motivation behind the above bounds is the case $E = B\cup C$. Another situation where quotients play a role appears in reversibility: Let Q be reversible. Then the equation of detailed balance $\pi(b)Q(b,c) = \pi(c)Q(c,b)$ for all $b\in B$ and $c \in C$ together with simple algebraic operations implies

$$\sum_{b\in B}\pi(b)Q(b,C) = \sum_{c\in C}\pi(c)Q(c,B)$$

or, equivalently,

$$\frac{\pi(B)}{\pi(C)} = \frac{\sum_{c\in C}\pi_C(c)Q(c,B)}{\sum_{b\in B}\pi_B(b)Q(b,C)}.$$

Therefore, the lower bound

$$\frac{\pi(B)}{\pi(C)} \geq \frac{\min_{c\in C} Q(c,B)}{\max_{b\in B} Q(b,C)} \tag{8}$$

follows immediately from the equality above. Note that the quotient $\pi(B)/\pi(C)$ may be rewritten as

$$\sum_{b\in B} \frac{1}{\sum_{c\in C} \frac{Q(b,c)}{Q(c,b)}}. \tag{9}$$

Hence, although no bounds are needed, a simple inspection shows that the lower bound (8) is simpler than the expression (9). If Q is not reversible then a way to "emulate" the procedure described above is the one given in the proof of Lemma 2.1.

(b) Note that, for any set D containing $B \cup C$ the quotients $\frac{\pi(B)}{\pi(C)}$ and $\frac{\pi_D(B)}{\pi_D(C)}$ are equal.

(c) Obviously, one is interested in the tightness of the bounds. Clearly, the precision of these bounds is non–decreasing in the ratios $\frac{\min_{c\in C} \quad Q(c,B)}{\max_{c\in C} \quad Q(c,B)}$, $\frac{\min_{b\in B} \quad Q(b,B)}{\max_{b\in B} \quad Q(b,B)}$, $\frac{\min_{c\in C} \quad Q(c,C)}{\max_{c\in C} \quad Q(c,C)}$, and $\frac{\min_{b\in B} \quad Q(b,C)}{\max_{b\in B} \quad Q(b,C)}$.

(d) The extension of the given results to Markov kernels on a finite state space E with a single recurrent set F and stationary distribution π is also straightforward: If the recurrent set F is known then we must replace the restriction $Q_{F\times F}$ of Q to the set F with Q itself throughout. If, however, F is unknown, (see, for example, the Markov kernel defined in (11)) then the given results are still valid, for subsets B, C such that $\pi(B)\cdot\pi(C) > 0$.

The following corollary establishes upper and lower bounds, for the restricted equilibrium probability of a subset of states. Part (b) is essentially a theorem due to Cronshaw and Kazarian (1997), p. 581. The proof of Parts (c) and (d) show how to (recursively) use the results (5) and (7) in order to improve these same bounds.

2.3 Corollary

(a) For all $\emptyset \neq B \subset C \subseteq E$, we have

$$\frac{L(B, C\setminus B)}{1 + L(B, C\setminus B)} \leq \pi_C(B) \leq \frac{U(B, C\setminus B)}{1 + U(B, C\setminus B)}.$$

(b) For all $\emptyset \neq B \subset E$, we have

$$\frac{L(B, B^c)}{1 + L(B, B^c)} \leq \pi(B) \leq \frac{U(B, B^c)}{1 + U(B, B^c)}.$$

(c) Let $\emptyset \neq B \subset E$. Moreover, let $\{B_0, B_1\}$ be a partition of B and $\{C_0, C_1\}$ be a partition of $\complement B$ such that $\min_{b\in B_0} Q(b,B) \leq \min_{b\in B_1} Q(b,B)$ and $\min_{c\in C_0} Q(c,B) \leq \min_{c\in C_1} Q(c,B)$. Then,

$$\pi(B) \geq \frac{[Q(C_0,B) + L_C \cdot Q(C_1,B)][1 + L_B]}{[1 + L_B - Q(B_0,B) - L_B \cdot Q(B_1,B)][1 + L_C] + [Q(C_0,B) + L_C \cdot Q(C_1,B)][1 + L_B]},$$

where $L_B := L(B_1,B_0)$, $L_C := L(C_1,C_0)$, $Q(B_0,B) := \min_{b\in B_0} Q(b,B)$, $Q(B_1,B) := \min_{b\in B_1} Q(b,B)$, $Q(C_0,B) := \min_{c\in C_0} Q(c,B)$ and $Q(C_1,B) := \min_{c\in C_1} Q(c,B)$.

(d) Let $\emptyset \neq B \subset E$. Moreover, let $\{B_0,B_1\}$ be a partition of B and $\{C_0,C_1\}$ be a partition of $\complement B$ such that $\min_{b\in B_0} Q(b,C) \leq \min_{b\in B_1} Q(b,C)$ and $\min_{c\in C_0} Q(c,C) \leq \min_{c\in C_1} Q(c,C)$. Then,

$$\pi(B) \leq \frac{[1+L_C - Q(C_0,C) - L_C\cdot Q(C_1,C)[1+L_B]}{[Q(B_0,B)+L_B Q(B_1,B)][1+L_C] + [1+L_C - Q(C_0,C) - L_C\cdot Q(C_1,C)][1+L_B]},$$

where $L_B := L(B_1,B_0)$, $L_C := L(C_1,C_0)$, $Q(B_0,C) := \min_{b\in B_0} Q(b,C)$, $Q(B_1,C) := \min_{b\in B_1} Q(b,C)$, $Q(C_0,C) := \min_{c\in C_0} Q(c,C)$ and $Q(C_1,C) := \min_{c\in C_1} Q(c,C)$.

Proof. Part (b) follows from Part (a) applied to the case $C = E$. The proof of Part (a) follows from (5) and (7) applied to the case B and $C = C\setminus B$ together with simple algebraic manipulations. We now prove Part (c). By (5) applied to $B = B_1$ and $C = B_0$, we obtain $\pi(B_1)/\pi(B_0) \geq L(B_1,B_0)$. Therefore,

$$(\pi_B(B_0),\pi_B(B_1)) \succeq (\frac{1}{1+L(B_1,B_0)}, \frac{L(B_1,B_0)}{1+L(B_1,B_0)}) =: \alpha_1.$$

Similarly,

$$(\pi_C(C_0),\pi_C(C_1)) \succeq (\frac{1}{1+L(C_1,C_0)}, \frac{L(C_1,C_0)}{1+L(C_1,C_0)}) =: \alpha_2.$$

These last two assertions together with the conditions $\min_{b\in B_0} Q(b,B) \leq \min_{b\in B_1} Q(b,B)$ and $\min_{c\in C_0} Q(c,B) \leq \min_{c\in B_1} Q(c,B)$ and (4) imply

$$\frac{\pi(B)}{\pi(\complement B)} \geq \frac{[\min_{c\in C_0} Q(c,B) + L(C_1,C_0)\min_{c\in C_1} Q(c,B)][1+L(B_1,B_0)]}{[1+L(B_1,B_0) - \min_{b\in B_0} Q(b,B) - L(B_1,B_0)\min_{b\in B_1} Q(b,B)][1+L(C_1,C_0)]}.$$

Our claim now follows from the inequality above together with simple algebraic manipulations. The proof of Part (d) is analogous to this one and we omit it. $\square$

Clearly, Parts (c) and (d) above may be extended to partitions of B and $\complement B$ with an arbitrary number of elements. However, the resulting bounds may not be simple.

3 First passage times

Let A be a nonempty proper subset of E. In this section, we define a new transition matrix. Its stationary distribution allows us to represent the mean first passage time to the set A as a quotient of the corresponding equilibrium probabilities of two events. We need first some preliminaries.

3.1 The first passage time to the set A

Let $X_n : (\Omega, P) \to E$, $n \in \mathbf{N}$, be an homogeneous Markov chain with transition matrix Q and initial distribution $P_{X_0} = \theta$. Let A be a nonempty subset of E and let $B := E\setminus A$. Assume that

(i) the pair (Q, A) satisfies the condition

($\star$) for all $t \in B$ there exists an integer $n_t \geq 1$ such that $Q^{n_t}(t, A) > 0$ and

(ii) $\theta(B) > 0$.

Clearly, if Q is irreducible then (i) is satisfied. The random variable $T_A : (\Omega, P) \to \mathbf{N} \cup \infty$ defined by

$$T_A := \inf\{n \in \mathbf{N}/X_n \in A\},$$

where $\inf \emptyset := \infty$, is called the first passage time to the set A. It is well known that, under condition ($\star$), we have $P[T_A < \infty] = 1$. Moreover, for $n \in \mathbf{N}$, we have

$$P^\theta[T_A > n] = \theta_{|B} Q^n_{|B\times B} \mathbf{1}.$$

Hence, the distribution of the random variable T_A satisfies

$$P^\theta[T_A = n] = P^\theta[T_A > n-1] - P^\theta[T_A > n] = \theta_{|B} Q^{n-1}_{|B\times B}(I - Q_{|B\times B})\mathbf{1}$$

if $n \geq 1$ and we have $P^\theta[T_A = 0] = \theta(A)$. It is well known that, under condition ($\star$), the inverse $(I - Q_{|B\times B})^{-1}$ exists and that

$$E^\theta[T_A] = \theta_{|B}(I - Q_{|B\times B})^{-1}\mathbf{1}, \tag{10}$$

cf. [6] Theorem 3.5.4(3).

The aim of the following results is to exhibit the mean first passage time to the set A as a quotient of two stationary probabilities of a suitable transition kernel. The following definition is a simple generalisation of the one given in [6], Theorem 3.2.2.

3.2 Definition

Let $Q \in \mathbf{R}^{E\times E}$, $\theta \in \mathbf{R}^E$ and $A \subset E$ as in 3.1. Let $Q^{\theta,A} \in \mathbf{R}^{E\times E}$ be the following transition matrix

$$Q^{\theta,A} = \begin{pmatrix} \mathbf{1}\theta_{|A} & \mathbf{1}\theta_{|B} \\ & \\ Q_{|B\times A} & Q_{|B\times B} \end{pmatrix}. \tag{11}$$

The matrix $Q^{\theta,A}$ is obtained from the matrix Q by putting all rows corresponding to the states in the set A to the same vector θ.

The Lemmas 3.3 and 3.5 are simple generalisations of [6], Theorem 6.2.2.

3.3 Lemma

Let the situation be as in Subsection 3.1 and consider Definition 3.2. Let $\mathbf{J}^\theta$ be the set of all states that are reachable from a state in the support S_θ of the distribution θ. The set $\mathbf{J}^\theta$ is the only recurrent set of the matrix $Q^{\theta,A}$ and at least one element of the set A belongs to $\mathbf{J}^\theta$.

Proof. Plainly, from a state $a \in A$ the chain can go only to states in $\mathbf{J}^\theta$ and, therefore, at least one element of the set A is also an element of $\mathbf{J}^\theta$. Moreover, Condition ($\star$) implies that the chain can go from any state to some state in the set A. Hence, all states in $E \setminus \mathbf{J}^\theta$ are transient. Since, from any state, the chain can reach A, and from there all states in S_θ and, hence, $\mathbf{J}^\theta$ we see that $\mathbf{J}^\theta$ is recurrent. Hence, the new matrix $Q^{\theta,A}$ has the single recurrent set $\mathbf{J}^\theta$, and at least one element of the set A is also an element of $\mathbf{J}^\theta$. $\square$

3.4 Notations and Remarks

(a) We denote by $\pi^{\theta,A} = (\pi^{\theta,A}_{|A}, \pi^{\theta,A}_{|B})$ the (unique) stationary distribution of the matrix $Q^{\theta,A}$. By Lemma 3.3, we have $\pi^{\theta,A}(A) > 0$.
(b) Let $B \subseteq E$. If $\theta(B) > 0$ then we have $\pi^{\theta,A}(B) > 0$.

3.5 Lemma

The stationary distribution $\pi^{\theta,A}$ for $Q^{\theta,A}$ satisfies the two equalities

$$\frac{\pi^{\theta,A}_{|B}}{\pi^{\theta,A}(A)} = \theta_{|B}(I - Q_{|B\times B})^{-1}, \tag{12}$$

and

$$\frac{\pi^{\theta,A}_{|A}}{\pi^{\theta,A}(A)} = \theta_{|B}(I - Q_{|B\times B})^{-1}Q_{|B\times A} + \theta_{|A}. \tag{13}$$

Proof. The equalities $Q^{\theta}_{|A\times A} = \mathbf{1}\theta_{|A}$ and $Q^{\theta}_{|A\times B} = \mathbf{1}\theta_{|B}$ show the following decomposition of the system $\pi^{\theta,A} = \pi^{\theta,A}Q^{\theta,A}$:

$$\begin{aligned}
\theta_{|A} + \frac{\pi^{\theta,A}_{|B}}{\pi^{\theta,A}(A)}Q_{|B\times A} &= \frac{\pi^{\theta,A}_{|A}}{\pi^{\theta,A}(A)},\\
\theta_{|B} + \frac{\pi^{\theta,A}_{|B}}{\pi^{\theta,A}(A)}Q_{|B\times B} &= \frac{\pi^{\theta,A}_{|B}}{\pi^{\theta,A}(A)}.
\end{aligned}$$

The second equality yields (12) and, by replacing the value of the vector $\frac{\pi^{\theta,A}_{|T}}{\pi^{\theta,A}(A)}$ in the first equation, we obtain (13). The lemma is now established. $\square$

It is now sufficient to consider formula (10) together with (12) to obtain a representation for the mean first passage time to the set A as a suitable quotient of stationary probabilities. This is subject of the next proposition.

3.6 Proposition

With the notations of Lemma 3.5, we have

$$E^{\theta}[T_A] = \frac{\pi^{\theta,A}(B)}{\pi^{\theta,A}(A)}.$$

3.7 Remarks

(a) We thus see that the single vector $\pi^{\theta,A}$ furnishes us with the mean first passage time to the set A. We will use this expression in Section 5 together with the results of Section 2 in order to investigate the mean first passage time to interesting subsets corresponding to a Markov chain describing the dynamics of an evolutionary game. As expected, the way how our theoretical results combine in order to achieve these results depends on this application.

(b) If the single recurrent set $\mathbf{J}^{\theta}$ is not known, the results of Section 2 are also applicable, see e.g. Remark 2.2(d).

4 An application: Evolutionary Game Theory

The model 4.2 below formalises the situation emerging in an evolutionary game. We sketch the situation briefly here. A more detailed description may be found in Kandori, Mailath and Rob (1993) (hereafter KMR). For related works on this theme see also Ellison (1993), Rhode and Stegeman (1996), Young (1993) and Canals and Vega–Redondo (1998).

4.1 Motivation (Kandori, Mailath and Rob (1993))

Consider the following 2×2 (symmetric) game

$$\begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\mathrm{II}} \\ & & s_1 & s_2 \\ \cline{3-4} \mathrm{I} & s_1 & a, a & g, c \\ \cline{3-4} & s_2 & c, g & d, d \\ \cline{3-4} \end{array} \quad , \tag{14}$$

where $S = \{s_1, s_2\}$ is the set of pure strategies and a, g, c and d are real numbers. If $a > c, d > g$ then the game is a <u>*coordination game*</u> and the set of pure strategy equilibria is given by $\{(s_1, s_1), (s_2, s_2)\}$. If $(a - c)(d - g) < 0$ then (14) is a game with a <u>*dominant strategy*</u>. KMR introduce the following discrete frame of an evolutionary game: Consider a group of N players each of which chooses his pure strategy, i.e., s_1 or s_2 at the beginning of each time period $t \in \mathbf{N}$. Next, the players are randomly matched and each pair plays the 2×2 coordination game (14). Let $x_t \in 0..N =: E$ be the number of players adopting strategy s_1 at period t. KMR assume that a player with strategy s_1 and s_2, respectively, gets a payoff

$$\pi_1(x_t) = \frac{x_t - 1}{N - 1} a + \frac{N - x_t}{N - 1} g \quad \text{and} \quad \pi_2(x_t) = \frac{x_t}{N - 1} c + \frac{N - x_t - 1}{N - 1} d, \tag{15}$$

if $x_t \in 1..(N-1)$. (Note that (15) corresponds to the expectation of the random payoff due to the random matching). In order to optimise its gain in the near future, a player chooses the best response against the current strategy distribution. This approach is formalised by means of any function $b : E \to E$ satisfying the following "Darwinian" property

$$(D) \qquad \operatorname{sign}(b(x) - x) = \operatorname{sign}(\pi_1(x) - \pi_2(x)), \qquad x \in 1..(N-1),$$

i.e., it is assume that the number of players choosing s_1 in the next period is given by

$$x_{t+1} = b(x_t). \tag{16}$$

In addition to this deterministic dynamics, KMR introduce the following random component: In each period each player "flips" its chosen strategy with a probability or <u>*mutation rate*</u> $\epsilon > 0$ (i.i.d across players and over time). These specifications yield the nonlinear stochastic difference equation

$$x_{t+1} = b(x_t) + Y_t(x_t) - W_t(x_t) = Y_t(x_t) + Z_t(x_t), \tag{17}$$

where $Y_t(x_t) : (\Omega, P) \to 0..N$ and $W_t(x_t) : (\Omega, P) \to 0..N$ are independent random variables having binomial distributions $Y_t(x_t) \sim \operatorname{Bin}(N - b(x_t), \epsilon)$ and $W_t(x_t) \sim \operatorname{Bin}(b(x_t), \epsilon)$ ($Z_t(x_t) \sim \operatorname{Bin}(b(x_t), 1 - \epsilon)$). Equation (17) defines an irreducible transition matrix whose stationary distribution describes the long run behaviour of this evolutionary game. The main interest of KMR was to analyse the asymptotic behaviour of the number of players playing each strategy as the mutation rate ϵ approaches 0, when the function b is given by 4.1(D).

A look at (17) shows that the Markov kernel describing the dynamics of an evolutionary game belongs to the class of transition matrices K, for which the kth row is the convolution of $\mathrm{Bin}(N-b(x),\epsilon)$ and $\mathrm{Bin}(b(x),1-\epsilon)$ and $b: E \to E$ is a function such that, from a certain value on, its graph is above the diagonal and below on the complement.

4.2 The model

Let $N \geq 2$ and let $E = 0..N$. Let $b: E \to E$ be a function on E and $\epsilon \in]0,1/2]$ and let $B_1 := \{x \in 1..(N-1) \ / \ b(x) < x\}$ and $B_2 := \{x \in 1..(N-1) \ / \ b(x) > x\}$. We assume that

(i) B_1 and B_2 are consecutive intervals,

(ii) $B_1 \cup B_2 = 1..(N-1)$, and

(iii) if $B_1 \neq \emptyset$ then $b(0) = 0$ and, if $B_2 \neq \emptyset$ then $b(N) = N$.

Note that there is no relationship between the adjustment b and the payoffs π_1 and π_2, here. This is the only information available on b. Moreover, if $b(0) = 0$ and $b(N) = N$ then redefine the sets B_i, $i \in 1..2$, as $B_1 := B_1 \cup \{0\}$ and $B_2 := B_2 \cup \{N\}$, if $b(0) > 0$ and $b(N) = N$ then $B_1 := \emptyset$ and $B_2 = E$, and, finally if $b(0) = 0$ and $b(N) < N$ then $B_1 := E$ and $B_2 = \emptyset$. We denote $b^* = \sup B_1$ ($\sup \emptyset := -\infty$) and $b_* = \inf B_2$ ($\inf \emptyset := \infty$). We call B_1 the *domain of attraction* of 0 and B_2 the domain of attraction of N. In light of our motivation, the domain of attraction of N (0) is the set of all possible initial number of players, for which the deterministic dynamic (16) eventually ends up with all players adopting strategy s_1 (s_2). Moreover, let $A^\epsilon = (A^\epsilon(0), A^\epsilon(1), \ldots, A^\epsilon(N)) : (\Omega, P) \to E^{N+1}$ and $B^\epsilon = (B^\epsilon(0), B^\epsilon(1), \ldots, B^\epsilon(N)) : (\Omega, P) \to E^{N+1}$ be independent random tuples having independent components and one dimensional binomial distributions given by

$$A^\epsilon(k) \sim \mathrm{Bin}(k,\epsilon) \qquad \text{and} \qquad B^\epsilon(k) \sim \mathrm{Bin}(k, 1-\epsilon), \qquad k \in E.$$

Set $S^\epsilon(k) := A^\epsilon(N-k) + B^\epsilon(k) : (\Omega, P) \to E$. Moreover, let $K^\epsilon \in \mathbf{R}^{E\times E}$ be the transition matrix defined by

$$K^\epsilon(k,\cdot) = P[S^\epsilon(b(k)) = \cdot], \qquad k \in E, \tag{18}$$

i.e., the distribution of the kth row of this Markov kernel is the convolution of the distributions $\mathrm{Bin}(N-b(k),\epsilon)$ and $\mathrm{Bin}(b(k), 1-\epsilon)$. We say that K^ϵ is generated by b. Note that since $\epsilon > 0$ all the elements in the matrix K^ϵ are strictly positive. It is well known then that the Markov kernel has a unique stationary distribution γ^ϵ. We also remark that all the elements belonging to a set $b^{-1}(y)$, $y \in E$, have the same transition probabilities, i.e., the Markov kernel K^ϵ is lumpable with respect to b. Let us denote by $X_n : (\Omega, P) \to E$, $n \in \mathbf{N}$, a Markov chain with initial distribution P_{X_0} and transition matrix K^ϵ. For all $B \subseteq E$, $B \neq \emptyset$, let $T_B : (\Omega, P) \to \mathbf{N} \cup \infty$ be the first passage time to the set B. Since K^ϵ is irreducible we have $P[T_B < \infty] = 1$.

Let $\tilde{b} : E \to E$ be given by $\tilde{b}(x) = N - b(N-x)$, $x \in E$. Then, the corresponding domains of attraction are $\tilde{B}_1 = 0..(N-b_*)$ and $\tilde{B}_2 = (N-b^*)..N$. Let us denote by $\tilde{K}^\epsilon \in \mathbf{R}^{E\times E}$ the Markov kernel generated by $\tilde{b}$, cf. (18). A moment of reflection shows $K^\epsilon(x,y) = \tilde{K}^\epsilon(N-x, N-y)$ and $\gamma^\epsilon(x) = \tilde{\gamma}^\epsilon(N-x)$, for all $x, y \in E$. Moreover, if we define, for all $n \geq 0$, $\tilde{X}_n := N - X_n : (\Omega, P) \to E$ then it is easy to see that, (i) $(\tilde{X}_n)_{n\geq 0}$ is a Markov chain with corresponding transition matrix $\tilde{K}^\epsilon$; and (ii) $E^{\tilde{\gamma}^\epsilon}[\tilde{T}_{y..N}] = E^{\gamma^\epsilon}[T_{0..(N-y)}]$ and $E^{\tilde{\gamma}^\epsilon}[\tilde{T}_{0..(y-1)}] = E^{\gamma^\epsilon}[T_{y..N}]$, for all $y \in 0..N$.

4.3 Remarks

(a) Let us now connect the game theoretical framework of KMR and the model 4.1. To be more precise, we investigate which 2×2 symmetric games (14) and which population sizes N define "Darwinian" adjustments bs satisfying 4.1(i)–(iii). Simple algebraic manipulations show that

(i) If $a - g - c + d = 0$ and $d - a + (g-d)N > 0$ then $B_1 = \emptyset$ and $B_2 = E$, i.e., $b_\star = 0$.

(ii) If $a - g - c + d = 0$ and $d - a + (g-d)N < 0$ then $B_1 = E$ and $B_2 = \emptyset$, i.e., $b^\star = N$ and $b_\star = \infty$.

Let now $a-c+d-g \neq 0$ and $\tilde{x} := \frac{N(d-g)+a-d}{a-c+d-g} \notin 1..(N-1)$. Assume first that $a-c+d-g > 0$.

(iii) If $\lfloor \tilde{x} \rfloor \leq 0$ then either $B_1 = \emptyset$ and $B_2 = E$ ($b_\star = 0$) or $B_1 = \{0\}$ and $B_2 = 1..(N-1)$ depending on the value of $b(0)$ ($b_\star = 1$).

(iv) If $1 \leq \lfloor \tilde{x} \rfloor \leq N-2$ then $B_1 = 0..\lfloor \tilde{x} \rfloor$ and $B_2 = \lceil \tilde{x} \rceil..N$, i.e., $b_\star = \lceil \tilde{x} \rceil$.

(v) If $\lfloor \tilde{x} \rfloor = N-1$ then either $B_1 = 0..(N-1)$ and $B_2 = \{N\}$ ($b_\star = N$) or $B_1 = E$ and $B_2 = \emptyset$ ($b^\star = N$ and $b_\star = \infty$) depending on $b(N) = N$ or $b(N) < N$.

(vi) If $\lfloor \tilde{x} \rfloor \geq N$ then $B_1 = E$ and $B_2 = \emptyset$, i.e., $b^\star = N$ and $b_\star = \infty$.

Let now $a - c + d - g < 0$.

(vii) If $\lfloor \tilde{x} \rfloor \leq 0$ then $B_1 = E$ and $B_2 = \emptyset$, i.e., $b_\star = \infty$.

(viii) If $\lfloor \tilde{x} \rfloor \geq N$ then $B_1 = \emptyset$ and $B_2 = E$, i.e., $b_\star = 0$.

Note that the cases (i)–(viii) subdivide an important subset of all pairs 2×2 symmetric games, population sizes N which define an evolutionary game as in 4.1. Thus, if (14) is the payoff matrix of a coordination matrix, i.e., $a > c$ and $d < g$ (and as a normalisation $a \geq d$) then we are dealing with the cases (iv)-(vi). It is also not difficult to see that if the game has as dominant strategy then we are dealing with one of the cases above. In the Appendix A we consider the case $\tilde{x} \in 1..(N-1)$.

(b) (The model of Rhode and Stegeman) As already noted, the main interest of KMR was to analyze the asymptotic behavior of the number of players playing each strategy as the mutation rate ϵ approaches 0, when the function b is given by 4.1(D). An improved and extended version of the results achieved by KMR is the Theorem on page 447 of Rhode and Stegeman (henceforth SR). As pointed out by these two authors, KMR either (i) do not restrict the values of $b(0)$ and $b(N)$ arguing that they do not affect their results or (ii) assume that if $\pi_1(N) < \pi_2(N)$,

$b(N) < N$, and if $\pi_1(0) > \pi_2(0)$, $b(0) > 0$, where $\pi_i(j)$, $i \in 1..2$, $j \in \{0, N\}$ are defined as in (15).
In the case (i), SR indicate that this fact is true if the game is a coordination game. But by two counterexamples, cf. [7], Footnote 4, they illustrate that the values of $b(0)$ and $b(N)$ are important for other kinds of games. In the case (ii) they show that this assumption may introduce anomalies, cf. [7], Footnote 1. Therefore, they assume that

$$b(0) = 0 \text{ and } b(N) = N. \tag{19}$$

Furthermore, they say that strategy s_1 is stable if $\pi_1(N-1) - \pi_2(N-1) > 0$, and strategy s_2 is stable if $\pi_1(1) - \pi_2(1) < 0$. Finally, because of (19), the model of SR implies $\{0\} \subseteq B_1 \neq \emptyset$, and $\{N\} \subseteq B_2 \neq \emptyset$.

(c) In order to obtain their results, both KMR and SR use a characterization of the stationary distribution due to Freidlin and Wentzel (1984) on the basis of a tree decomposition. Using the techniques developed in Sections 2 and 3, we are able to analyze in more detail the long run behavior of the Markov chain given in 4.2 as the mutation rate approaches 0 and, therefore, by Remark (b) we are also able to analyze in more detail the dynamics of an evolutionary game.

(d) Although the model 4.2 allows more general payoffs than (15), such as the one given by

$$\pi_1(x) = p_1(x)a + (1 - p_1(x))g \text{ and } \pi_2(x) = p_2(x)c + (1 - p_2(x))d, \qquad x \in 1..(N-1),$$

it is not a generalization or extension of the models given by KMR and SR. The reason is that the Darwinian adjustment 4.1(D) and the properties 4.2(i)–(iii) of the triple (b, B_1, B_2) are equivalent: both of them depend only on the value $\tilde{x}$ for which $b(\tilde{x}) = \tilde{x}$. We could also write on the right hand side of 4.1(D) $\operatorname{sign}(\pi_1(x) - \pi_2(x))$, where this time π_i, $i \in 1..2$, is given above. The difference between the three models are the values of $b(0)$ and $b(N)$: see the Remark (b) above together with the fact that, contrary to the model of RS, the values of $b(0)$ and (or) $b(N)$ are only fixed when $B_1 = \emptyset$ and (or) $B_2 = \emptyset$. To be more precise, in 4.2 we have that $b(0) = 0$ if $\pi_1(1) - \pi_2(1) < 0$, and $b(N) = N$ if $\pi_1(N-1) - \pi_2(N-1) > 0$.

For the purposes of analysis we may assume without loss of generality that $B_2 \neq \emptyset$.

5 Asymptotics for the stationary distribution

For the proof of the main result of this and the next section we need first some definitions and lemmas. Since we are interested in the asymptotic behavior of various parameters of the Markov chain $(X_n)_{n \geq 0}$ as ϵ decreases to 0, we will use the "big $\mathcal{O}$" notation in the following sense: We write $f(x) = \mathcal{O}(g(x))$ $(x \to 0)$ (or shorter $f = \mathcal{O}(g)$), if there exist constants $C, c > 0$ such that $c|g(x)| \leq f(x) \leq C|g(x)|$, for x sufficiently small. Moreover, we write $f \leq \mathcal{O}(g)$ if there exists a constant $C > 0$ such that $|f(x)| \leq C|g(x)|$ for x sufficiently small and we write $\mathcal{O}(g) \leq f$ if there exists a constant $C > 0$ such that $C|g(x)| \leq |f(x)|$ for x sufficiently small. The following rules are plain: Let $a_1, a_2 \in \mathbf{N}$. Then $\frac{\mathcal{O}(\epsilon^{a_1})}{\mathcal{O}(\epsilon^{a_2})} = \mathcal{O}(\epsilon^{a_1 - a_2})$, $\mathcal{O}(\epsilon^{a_1}) + \mathcal{O}(\epsilon^{a_2}) = \mathcal{O}(\epsilon^{\min\{a_1, a_2\}})$ and $\mathcal{O}(\epsilon^{a_1}) \cdot \mathcal{O}(\epsilon^{a_2}) = \mathcal{O}(\epsilon^{a_1 + a_2})$. The generalization to an arbitrary number of a_i's is straightforward.

Assertion (20) of the following result is Lemma 2 of KMR. We extend it for our purposes.

5.1 Lemma

Let $0 \leq x, y \leq N$. We have

$$P[S^\epsilon(y) = x] = \mathcal{O}(\epsilon^{|y-x|}). \tag{20}$$

Moreover,

$$P[S^\epsilon(y) \geq x] = \mathcal{O}(\epsilon^{x-y}), \quad y \leq x, \quad P[S^\epsilon(y) \leq x] = \mathcal{O}(\epsilon^{y-x}), \quad y \geq x, \tag{21}$$

and

$$P[S^\epsilon(y) \leq x] = \mathcal{O}(1), \quad y \leq x, \quad P[S^\epsilon(y) \geq x] = \mathcal{O}(1), \quad y \geq x. \tag{22}$$

Proof. By symmetry, it is enough to prove (20) and the right hand side of (21), for the case $0 \leq x \leq y \leq N$. The probability $P[S^\epsilon(y) = x]$ may be written as

$$\sum_{\substack{0 \leq j_1 \leq N-y \\ 0 \leq j_2 \leq y \\ j_1 + j_2 = x}} \binom{N-y}{j_1} \binom{y}{j_2} \epsilon^{j_1 - j_2 + y} (1-\epsilon)^{N-y-j_1+j_2}.$$

Since the smallest power with which an ϵ appears in the expression above is $y - x$, namely when $j_1 = 0$ and $j_2 = x$, we have $P[S^\epsilon(y) = x] = \mathcal{O}(\epsilon^{y-x})$. Therefore,

$$P[S^\epsilon(y) \leq x] = \sum_{z=0}^{x} P[S^\epsilon(y) = z] = \sum_{z=0}^{x} \mathcal{O}(\epsilon^{y-z}) = \mathcal{O}(\epsilon^{y-x}).$$

The proof of (22) is analogue to this one. $\square$

The next proposition is one of the building blocks used to prove the results concerning the asymptotics of the stationary distribution and the first passage times. Its proof, among other things, exemplifies how to make use of the bounds given in Proposition 2.1 in order to obtain meaningful statements.

5.2 Proposition

Let $b : E \to E$ be as in 4.2 and let $Q^\epsilon \in \mathbf{R}^{E \times E}$ be an irreducible Markov kernel with unique stationary distribution π^ϵ.

(a) If,

(i) for all $x \in b_*..(N-1)$, we have $Q^\epsilon(x, \cdot\,) = P[S^\epsilon(b(x)) = \cdot\,]$ and

(ii) for all $y \in E$, we have $Q^\epsilon(N, y) = \mathcal{O}(\epsilon^{N-y})$

then, for all $m \in B_2$ and for all $x \in m..N$, we have $\pi^\epsilon_{m..N}(x) = \mathcal{O}(\epsilon^{N-x})$.

(b) If,

(i) for all $x \in 1..(b^*)$, we have $Q^\epsilon(x, \cdot\,) = P[S^\epsilon(b(x)) = \cdot\,]$ and

(ii) for all $y \in E$, we have $Q^\epsilon(0, y) = \mathcal{O}(\epsilon^y)$

then, for all $m \in B_1$ and for all $x \in 0..m$, we have $\pi^\epsilon_{0..m}(x) = \mathcal{O}(\epsilon^x)$.

Proof. (a) If $B_2 = \{N\}$ then the claim is trivial. Suppose now that $(0 \leq) b_* \leq N-1$. We first prove

$$\frac{\pi^\epsilon((x+1)..N)}{\pi^\epsilon(x)} = \mathcal{O}(\epsilon^{N-x}), \tag{23}$$

and

$$\pi^\epsilon_{x..N}(y) = \mathcal{O}(\epsilon^{N-y}), \qquad y \in x..N, \tag{24}$$

for all $x \in b_*..(N-1)$. For this sake, we use (backward) mathematical induction. By (5) and (7) applied to the case $B = \{N\}$ and $C = \{N-1\}$, we obtain

$$\frac{Q^\epsilon(N-1,N)}{1-Q^\epsilon(N,N)} \leq \frac{\pi^\epsilon(N)}{\pi^\epsilon(N-1)} \leq \frac{1-Q^\epsilon(N-1,N-1)}{Q^\epsilon(N,N-1)}. \tag{25}$$

By the assumptions (i) and (ii), we have

$$\frac{1-Q^\epsilon(N,N)}{Q^\epsilon(N-1,N)} = \frac{Q^\epsilon(N,0..(N-1))}{Q^\epsilon(N-1,N)} = \frac{\mathcal{O}(\epsilon)}{P[S^\epsilon(b(N-1)) = N]} = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)} = \mathcal{O}(\epsilon),$$

where the last equality follows from $b(N-1) = N$ together with (20). Similarly,

$$\frac{Q^\epsilon(N,N-1)}{1-Q^\epsilon(N-1,N-1)} = \frac{Q^\epsilon(N,N-1)}{Q^\epsilon(N-1,N)+Q^\epsilon(N-1,0..(N-2))} = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)+\mathcal{O}(\epsilon^2)} = \mathcal{O}(\epsilon).$$

These two facts together with (25) imply $\frac{\pi^\epsilon(N-1)}{\pi^\epsilon(N)} = \mathcal{O}(\epsilon)$. Hence, there exist constants $C, c > 0$ such that, for ϵ sufficiently small, we have $c\epsilon \leq \frac{\pi^\epsilon(N-1)}{\pi^\epsilon(N)} \leq C\epsilon$. Therefore,

$$\left(\frac{c\epsilon}{c\epsilon+1}, \frac{1}{c\epsilon+1}\right) \preceq (\pi^\epsilon_{(N-1)..N}(N-1), \pi^\epsilon_{(N-1)..N}(N)) \preceq \left(\frac{C\epsilon}{C\epsilon+1}, \frac{1}{C\epsilon+1}\right).$$

Clearly, this claim implies $\pi^\epsilon_{(N-1)..N}(N-1) = \mathcal{O}(\epsilon)$ and $\pi^\epsilon_{(N-1)..N}(N) = \mathcal{O}(1)$. This is (23) and (24), for $N-1$. Suppose now that the claim is already proved, for all $x \geq b_*+1$. By (3) and (6) applied to the case $B = x..N$ and $C = x-1$, we obtain

$$\frac{Q^\epsilon(x-1,x..N)}{1-\sum_{y=x}^N \pi^\epsilon_{x..N}(y)Q^\epsilon(y,x..N)} \leq \frac{\pi^\epsilon(x..N)}{\pi^\epsilon(x-1)} \leq \frac{1-Q^\epsilon(x-1,x-1)}{\sum_{y=x}^N \pi^\epsilon_{x..N}(y)Q^\epsilon(y,x-1)}. \tag{26}$$

By (20) together with the inequality $b(y) \geq y$ (then $y \geq x \geq b_*+1$), (21), the assumptions (i) and (ii), the fact $b(x-1) \in x..N$ and the induction hypothesis, we have

$$\begin{aligned}
&\frac{\sum_{y=x}^N \pi^\epsilon_{x..N}(y)Q^\epsilon(y,x-1)}{1-Q^\epsilon(x-1,x-1)} \\
=\;& \frac{\sum_{y=x}^{N-1} \pi^\epsilon_{x..N}(y)Q^\epsilon(y,x-1) + \pi^\epsilon_{x..N}(N)Q^\epsilon(N,x-1)}{Q^\epsilon(x-1,0..(x-2)) + Q^\epsilon(x-1,x..N)} \\
=\;& \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{N-y})\cdot\mathcal{O}(\epsilon^{b(y)-x+1}) + \mathcal{O}(1)\cdot\mathcal{O}(\epsilon^{N-x+1})}{\mathcal{O}(1)} \\
=\;& \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{N-y+b(y)-x+1}) + \mathcal{O}(\epsilon^{N-x+1})}{\mathcal{O}(1)} = \mathcal{O}(\epsilon^{N-x+1}).
\end{aligned}$$

An analogous argument shows

$$\begin{aligned}
&\frac{1-\sum_{y=x}^{N}\pi^{\epsilon}_{x..N}(y)Q^{\epsilon}(y,x..N)}{Q^{\epsilon}(x-1,x..N)}\\
=\;&\frac{\sum_{y=x}^{N}\pi^{\epsilon}_{x..N}(y)Q^{\epsilon}(y,0..(x-1))}{Q^{\epsilon}(x-1,x..N)}\\
=\;&\frac{\sum_{y=x}^{N}\mathcal{O}(\epsilon^{N-y+b(y)-x+1})}{\mathcal{O}(1)}=\mathcal{O}(\epsilon^{N-x+1}).
\end{aligned}$$

Therefore, from these two last assertions together with (26), it follows

$$\frac{\pi^{\epsilon}(x..N)}{\pi^{\epsilon}(x-1)}=\mathcal{O}(\epsilon^{N-x+1}).$$

Similarly as in the case $x=N-1$, this claim implies $\pi^{\epsilon}_{(x-1)..N}(x-1)=\mathcal{O}(\epsilon^{N-x+1})$ and $\pi^{\epsilon}_{(x-1)..N}(x..N)=\mathcal{O}(1)$. Since the probability vector $\pi^{\epsilon}_{(x-1)..N}$ may be written as

$$\pi^{\epsilon}_{(x-1)..N}=\pi^{\epsilon}_{(x-1)..N}(x-1)\delta_{x-1}+\pi^{\epsilon}_{(x-1)..N}(x..N)\pi^{\epsilon}_{x..N},$$

the claim (24), for $x-1$ follows from the induction hypothesis.

(b) Define $\tilde{Q}^{\epsilon}\in\mathbf{R}^{E\times E}$ as $\tilde{Q}^{\epsilon}(x,y):=Q^{\epsilon}(N-x,N-y)$, $x,y\in E$. A moment of reflection shows that $\tilde{Q}^{\epsilon}$ is an irreducible Markov kernel with stationary distribution $\tilde{\pi}^{\epsilon}$ given by $\tilde{\pi}^{\epsilon}(x)=\pi(N-x)$, $x\in E$. Moreover, $\tilde{Q}^{\epsilon}$, the function $\tilde{b}$, cf. the last paragraph in 4.2, and $\tilde{B}_2:=(N-b^*)..N$ satisfy the conditions of Part (a). Therefore, for all $x\in(N-b^*)..N$, we obtain $\tilde{\pi}^{\epsilon}_{x..N}(y)=\mathcal{O}(\epsilon^{N-y})$ or equivalently $\pi^{\epsilon}_{0..x}(y)=\mathcal{O}(\epsilon^{y})$, for all $y\in 0..x$ and, for all $x\in B_1$. □

5.3 Remarks

(a) Note that, in 5.2(a), we may write f instead of b_*, for each $f\in b_*..(N-1)$. Similarly, in 5.2(b), we may write f instead of b^*, for each $f\in 1..b^*$.
(b) By (20) together with the equality $b(N)=N$, we have $K^{\epsilon}(N,y)=P[S^{\epsilon}(N)=y]=\mathcal{O}(\epsilon^{N-y})$, $y\in E$. Hence, the Markov kernel K^{ϵ} satisfies the conditions (a)(i)–(ii) of the lemma above. Moreover, if $B_1\neq\emptyset$ then a similar argumentation shows $K^{\epsilon}(0,y)=P[S^{\epsilon}(0)=y]=\mathcal{O}(\epsilon^{y})$, $y\in E$. Hence, K^{ϵ} satisfies conditions (b)(i)–(ii).

With these results at hand, we are now able to state and prove the main result of this section.

5.4 Theorem

(a) Suppose that $B_1,B_2\neq\emptyset$. For the speed of convergence of the various elements of the stationary distribution γ^{ϵ}, we have

(i) For all $x\in B_2$, we have $\gamma^{\epsilon}_{B_2}(x)=\mathcal{O}(\epsilon^{N-x})$ and, for all $y\in B_1$, we have $\gamma^{\epsilon}_{B_1}(y)=\mathcal{O}(\epsilon^{y})$. Moreover, $\frac{\gamma^{\epsilon}(B_1)}{\gamma^{\epsilon}(B_2)}=\mathcal{O}(\epsilon^{N-2b_*+1})$.

(ii) If $b_*<\frac{N+1}{2}$ then, for all $x\in B_2$ we have $\gamma^{\epsilon}(x)=\mathcal{O}(\epsilon^{N-x})$ and for all $y\in B_1$, we have $\gamma^{\epsilon}(y)=\mathcal{O}(\epsilon^{y+N-2b_*+1})$. Moreover, $\gamma^{\epsilon}(B_1)=\mathcal{O}(\epsilon^{N-2b_*+1})$ and $\gamma^{\epsilon}(B_2)=\mathcal{O}(1)$.

(iii) If $b_* = \frac{N+1}{2}$ then, for all $x \in B_2$ we have $\gamma^\epsilon(x) = \mathcal{O}(\epsilon^{N-x})$ and for all $y \in B_1$, we have $\gamma^\epsilon(y) = \mathcal{O}(\epsilon^y)$. Moreover, $\gamma^\epsilon(B_1) = \mathcal{O}(1)$ and $\gamma^\epsilon(B_2) = \mathcal{O}(1)$.

(iv) If $b_* > \frac{N+1}{2}$ then, for all $x \in B_2$, we have $\gamma^\epsilon(x) = \mathcal{O}(\epsilon^{2b_*-x-1})$ and, for all $y \in B_1$, we have $\gamma^\epsilon(y) = \mathcal{O}(\epsilon^y)$. Moreover, $\gamma^\epsilon(B_1) = \mathcal{O}(1)$ and $\gamma^\epsilon(B_2) = \mathcal{O}(\epsilon^{2b_*-N-1})$.

(b) If $B_1 = \emptyset$ then $\gamma^\epsilon(x) = \mathcal{O}(\epsilon^{N-x})$, for all $x \in E$.

(c) If, otherwise, $B_2 = \emptyset$ then $\gamma^\epsilon(x) = \mathcal{O}(\epsilon^x)$, for all $x \in E$.

Proof. The first two claims of Item (a) (i) and Part (b) follow immediately from Proposition 5.2 together with (18). For the proof of the last claim of (a) (i), let in (3) $Q = K^\epsilon$, $B = B_1$, and $C = B_2$. We obtain

$$\frac{\gamma^\epsilon(B_1)}{\gamma^\epsilon(B_2)} = \frac{\sum_{y \in B_2} \gamma^\epsilon_{B_2}(y) P[S^\epsilon(b(y)) \in B_1]}{\sum_{x \in B_1} \gamma^\epsilon_{B_1}(x) P[S^\epsilon(b(x)) \in B_2]}. \tag{27}$$

Moreover, by (21), we have $P[S^\epsilon(b(y)) \in B_1] = \mathcal{O}(\epsilon^{b(y)-b^*})$, $y \in B_2$, and $P[S^\epsilon(b(x)) \in B_2] = \mathcal{O}(\epsilon^{b_*-b(x)})$, $x \in B_1$. These facts together with (27) and our first two claims imply

$$\frac{\gamma^\epsilon(B_1)}{\gamma^\epsilon(B_2)} = \frac{\sum_{y \in B_2} \mathcal{O}(\epsilon^{N-y}) \cdot \mathcal{O}(\epsilon^{b(y)-b^*})}{\sum_{x \in B_1} \mathcal{O}(\epsilon^x) \cdot \mathcal{O}(\epsilon^{b_*-b(x)})} = \frac{\mathcal{O}(\epsilon^{N-b^*})}{\mathcal{O}(\epsilon^{b_*})} = \mathcal{O}(\epsilon^{N-2b_*+1}). \tag{28}$$

This is the claim.

In view of (ii), let $b_* < \frac{N+1}{2}$. From (28) follows $\gamma^\epsilon(B_2) = \mathcal{O}(1)$ and $\gamma^\epsilon(B_1) = \mathcal{O}(\epsilon^{N-2b_*+1})$. Moreover, by (i) we have

$$\gamma^\epsilon(x) = \gamma^\epsilon(B_2) \cdot \mathcal{O}(\epsilon^{N-x}) = \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^{N-x}) = \mathcal{O}(\epsilon^{N-x}),$$

for all $x \in B_2$ and

$$\gamma^\epsilon(x) = \gamma^\epsilon(B_1) \cdot \mathcal{O}(\epsilon^x) = \mathcal{O}(\epsilon^{N-2b_*+1}) \cdot \mathcal{O}(\epsilon^x) = \mathcal{O}(\epsilon^{N-2b_*+1+x}),$$

for all $x \in B_1$.

We next prove (iii). If $b_* = \frac{N+1}{2}$ then by (28) we have $\gamma^\epsilon(B_2) = \mathcal{O}(1)$ and $\gamma^\epsilon(B_1) = \mathcal{O}(1)$ and, hence,

$$\gamma^\epsilon(x) = \gamma^\epsilon(B_2) \cdot \gamma^\epsilon_{B_2}(x) = \gamma^\epsilon(B_2) \cdot \mathcal{O}(\epsilon^{N-x}) = \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^{N-x}) = \mathcal{O}(\epsilon^{N-x}),$$

for all $x \in B_2$ and

$$\gamma^\epsilon(x) = \gamma^\epsilon(B_1) \cdot \gamma^\epsilon_{B_1}(x) = \gamma^\epsilon(B_1) \cdot \mathcal{O}(\epsilon^x) = \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^x) = \mathcal{O}(\epsilon^x),$$

for all $x \in B_1$. The proof of the Assertion (iv) is similar. $\square$

5.5 Remark

If $\{0,1\} \subseteq B_1$ or $\{N-1,N\} \subseteq B_2$ then the proof of [7], Theorem on page 447, yields implicitly some speeds of convergence for the stationary distribution γ^ϵ. The aim of this paragraph is to recover them. Let first $\{N-1,N\} \subseteq B_2$. For each $x \in E$, let us denote by v_x the speed of convergence of the corresponding x–tree. Then by the proof of this result we have that if $B_1 \neq \emptyset$ then

$$v_N = b_\star,\ v_z > b_\star,\ z \in b_\star..(N-1),\ v_0 = N - b^\star,\quad \text{and}\quad v_z > N - b^\star,\ z \in 1..b^\star. \tag{29}$$

It is easy to see that the speed of convergence of the sum over all z–trees, $z \in B_2$, equals $b_\star$, whereas the speed of convergence of the sum over all z–trees, $z \in B_1$, is $N - b^\star$. That is we have $v_{B_1} = N - b^\star$ and $v_{B_2} = b_\star$. Using the characterisation of a stationary distribution by means of z–trees we may write the quotient of two equilibrium probabilities as the quotient of the corresponding sums over all the probabilities of the z–trees. This fact together with the speeds of convergence derived above, and standard operations on order of convergence implies

(i)

$$\frac{\gamma^\epsilon(B_1)}{\gamma^\epsilon(B_2)} = \frac{\mathcal{O}(\epsilon^{b_\star})}{\mathcal{O}(\epsilon^{N-b^\star})} = \mathcal{O}(\epsilon^{2b_\star - N - 1}),$$

i.e., the second claim in 5.4(i).

(ii) Let $b_* < \frac{N+1}{2}$. By (i) we have $\gamma^\epsilon(B_1) = \mathcal{O}(\epsilon^{N+1-2b_\star})$ and $\gamma^\epsilon(B_2) = \mathcal{O}(1)$, i.e., the second claim of 5.4(ii). Further, using the facts derived in the first paragraph we obtain

$$\gamma^\epsilon(N) = \frac{\gamma^\epsilon(N)}{\gamma^\epsilon(B_2)} \cdot \gamma^\epsilon(B_2) = \frac{\mathcal{O}(\epsilon^{v_N})}{\mathcal{O}(\epsilon^{v_{B_2}})} \cdot \gamma^\epsilon(B_2) = \frac{\mathcal{O}(\epsilon^{b_\star})}{\mathcal{O}(\epsilon^{b_\star})} \cdot \mathcal{O}(1) = \mathcal{O}(1),$$

and if $x \in B_2 \setminus \{N\}$, there exists a natural number $\alpha_x \geq b_\star + 1$ such that

$$\gamma^\epsilon(x) = \mathcal{O}(\epsilon^{\alpha_x - b_\star}).$$

Claim which is implied by 5.4(ii). Using similar arguments one can show that

$$\gamma^\epsilon_{B_2}(x) = \mathcal{O}(\epsilon^{\alpha_x}).$$

Claim which is implied by 5.4(i). Analogously, we obtain

$$\gamma^\epsilon(0) = \mathcal{O}(\epsilon^{N-2b_\star+1}),$$

and for all $y \in B_1 \setminus \{0\}$ there exists a natural number $\beta_y \geq N - b^\star + 1$ such that

$$\gamma^\epsilon(y) = \mathcal{O}(\epsilon^{N-\beta_y-b_\star}) \qquad \text{and} \qquad \gamma^\epsilon_{B_1}(y) = \mathcal{O}(\epsilon^{b_\star - \beta_y - 1}).$$

Claims which are implied by 5.4(i)–(ii).

(iii) Let now $b_* = \frac{N+1}{2}$. By Item (i) above we obtain $\gamma^\epsilon(B_1) = \gamma^\epsilon(B_2) = \mathcal{O}$. By (29) we have $v_N = v_0 = \frac{N+1}{2}$, and $v_x > \frac{N+1}{2}$ for all $x \in 1..(N-1)$. Therefore,

$$\gamma^\epsilon(N) = \frac{\gamma^\epsilon(N)}{\gamma^\epsilon(B_2)} \cdot \gamma^\epsilon(B_2) = \frac{\mathcal{O}(\epsilon^{v_N})}{\mathcal{O}(\epsilon^{v_{B_2}})} \cdot \gamma^\epsilon(B_2) = \mathcal{O}(1).$$

Similarly, we have $\gamma^\epsilon(1) = \mathcal{O}(1)$. Furthermore, for all $x \in E \setminus \{0, 1N-1\}$ we have

$$\gamma^\epsilon(x) = \frac{\gamma^\epsilon(x)}{\gamma^\epsilon(B_1)} \cdot \gamma^\epsilon(B_1) = \frac{\mathcal{O}(\epsilon^{v_x})}{\mathcal{O}(\epsilon^{v_{B_1}})} \cdot \mathcal{O}(1) = \mathcal{O}(\epsilon^{v_x - \frac{N+1}{2}}) = \mathcal{O}(\epsilon^{\alpha_x}),$$

for an $\alpha_x \geq 1$. Claim which is implied by 5.4(iii).

Analogous results hold if $B_1 \neq \emptyset$, $N-1..N \subseteq B_2$ and $b_\star > \frac{N+1}{2}$, and if $B_2 \neq \emptyset$ and $0..1 \subseteq B_1$. Therefore, Theorem 5.4 extends the speeds of convergence given by the proof of this theorem. Furthermore, Theorem 5.4 is necessary for the computation of the order of convergence of the mean first passage time to certain subsets A of $1..N$, when the initial distribution of the Markov chain describing the dynamics of the evolutionary game 4.1 is the corresponding stationary distribution.

Taking into account Theorem 5.4 and Remark 4.3(a) we may compute the speed of convergence of the entries of the stationary distribution of a Markov chain emerging in the evolutionary games considered in 4.3(a) as the mutation rate approaches 0. As a corollary, we obtain the corresponding limit distributions $\lim_{\epsilon\to 0}\gamma^\epsilon$. We compile them in the following

5.6 Corollary

Consider the Remark 4.3(a).

(a) If $a - g - c + d = 0$ and $d - a + (g-d)N > 0$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(N) = 1$, whereas if $a - g - c + d = 0$ and $d - a + (g-d)N < 0$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(0) = 1$.

Let now $a - g - c + d > 0$.

(b) If $\lfloor \tilde{x} \rfloor < \frac{N-1}{2}$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(N) = 1$. If $\lfloor \tilde{x} \rfloor = \frac{N-1}{2}$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(0) = \lim_{\epsilon\to 0}\gamma^\epsilon(N) = \frac{1}{2}$. Finally, if $\lfloor \tilde{x} \rfloor > \frac{N-1}{2}$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(0) = 1$.

Let finally $a - g - c + d < 0$.

(c) If $\lfloor \tilde{x} \rfloor \leq 0$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(0) = 1$, whereas if $\lfloor \tilde{x} \rfloor \geq N$ then $\lim_{\epsilon\to 0}\gamma^\epsilon(N) = 1$.

5.7 Remarks

(a) For the case $B_1, B_2 \neq \emptyset$, the limit results given above follow from [7], Part(i) of the Theorem on page 447. In order to see this assertion, note that if, for example, $\lfloor \tilde{x} \rfloor < \frac{N-1}{2}$ then $\pi_1(\frac{N}{2}) - \pi_2(\frac{N}{2}) > 0$ or $\pi_1(\frac{N}{2}) - \pi_2(\frac{N}{2}) < 0$ according to whether N is even or odd. Therefore, the Claim 5.6(b) follows from this Theorem. The Claims 5.6(b) and (c) follow in a similar way.

(b) Note that the limit in the case 5.6(a) the depends on both, the entries of the payoff matrix (16) and the population size N. This fact is not new, cf. [7], Section 2.

Both, Proposition 5.2 and Theorem 5.4, are somehow amazing if we take into account that, although we do not know in detail the function b first, we were capable to compute the speed of convergence of the entries of the stationary distribution as the mutation rate ϵ approaches

0 and second, these parameters do not depend on b. Results 5.4(ii)–(iv) show that the position of the point b_* or, equivalently, the (relative) sizes of the domains of attraction are decisive: If b_* is smaller than $\frac{N+1}{2}$ or, equivalently, if the domain of attraction of N is larger than that of 0 then, for ϵ sufficiently small, the whole mass of the stationary distribution is located in B_2 and, in particular, at N. Moreover, for low values of the mutation rate ϵ, the speed of convergence at which the stationary probability of the other members of the set B_2 approaches 0 is inversely proportional to the magnitude of the element. This holds, irrespective of the precise nature of b which may be highly asymmetric in how it behaves in this set. The rate of convergence at which the stationary probability of the set B_1 approaches 0 also depends on the size of B_2, only. A smaller b_* implies a faster convergence of $\gamma^\epsilon(B_1)$ to 0. The speed of convergence to 0 of $\gamma^\epsilon(x)$ for an element $x \in B_1$ is exponential with x; the proportionality factor is $\gamma^\epsilon(B_1)$. Hence, restricted to the set B_1, the element 0 is the one whose stationary probability approaches 0 at the lowest rate. If, otherwise, $b_* = \frac{N+1}{2}$ then the limit distribution is symmetric, i.e., each of the sets B_1 and B_2 absorbs half of the whole mass. Moreover, within these sets the states 0 and N are the favorites. The case $b_* > \frac{N+1}{2}$ is just a relabeling of the case $b_* < \frac{N+1}{2}$. The speed of convergence of the various elements of the conditional stationary distributions $\gamma^\epsilon_{B_1}$ ($\gamma^\epsilon_{B_2}$) do not depend on (i) the exact knowledge of the function b and (ii) the relative sizes of the domains of attraction. Conditional on being on the set B_1 (B_2), the state 0 absorbs the whole mass, whereas the speed of convergence at which $\gamma^\epsilon(x)$, $x \in B_1 \setminus \{0\}$ ($x \in B_2 \setminus \{N\}$), approaches 0 ($N$) is proportional (inversely proportional) to the magnitude of the element.

As a final remark, we point out that similar results for the limits $\lim_{\epsilon\to 0} \gamma^\epsilon(x)$, $x \in E$, were already stated by KMR, cf. [5], Section 8.

6 Asymptotics for the first passage time

The aim of this section is to apply our given theoretical results to investigate the asymptotic behaviour of the mean first passage time to various interesting subsets, when the initial distribution of the Markov chain $(X_n)_{n\geq 0}$ is the corresponding stationary distribution γ^ϵ. This scenario describes the dynamics of an evolutionary game that has been ongoing for a large period of time. The result achieved by KMR is the following

6.1 Theorem (Kandori, Mailath and Rob (1993))

Consider the stage game (14) and let $\tilde{x} := \frac{N(d-b)+a-d}{a-c+d-b} \neq \frac{N}{2}$. For any population size $N \in 2, 4, 8, \ldots$ and for the adjustment process $b^D : E \to E$ given by

$$b^D(x) = \begin{cases} 0 & , \quad x < \tilde{x}, \\ N & , \quad x > \tilde{x}, \end{cases}$$

we have

(a) if $\tilde{x} < \frac{N-2}{3}$ then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = 0$ and

(b) if $\tilde{x} > \frac{N+1}{3}$ then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$,

where $B_2 = \lceil \tilde{x} \rceil..N$.

The results of Sections 2 and 3 enable to go far beyond this result since they allow the treatment of any function $b : E \to E$ satisfying solely the conditions 4.2(i)–(iii). Moreover, Theorem 6.1 does not provide information on the asymptotics of $E^{\gamma^\epsilon}[T_{B_2}]$ when $\tilde{x} \in \frac{N-2}{3}..\frac{N+1}{3}$ or, equivalently, when $\lceil \tilde{x} \rceil = \lceil \frac{N+1}{3} \rceil$. (Note that $\tilde{x} < \frac{N-2}{3}$ if and only if $\lceil \tilde{x} \rceil \le \lceil \frac{N-2}{3} \rceil$ and $\tilde{x} > \frac{N+1}{3}$ implies $\lceil \tilde{x} \rceil \ge \lceil \frac{N+4}{3} \rceil$.)

Although the proof of the next theorem is long and technical, we find it instructive. It shows, among other things, that in order to achieve useful building blocks or bounds (3) and (6), for the proof of the final result, one must first take a look at the transition matrix in order to choose appropriate sets B and C. Moreover, the following theorem fills the gap described above.

6.2 Theorem

Let $b : E \to E$ satisfy 4.2(i)–(iii).

(a) If $b_* = 0$ then $E^{\gamma^\epsilon}[T_N] = \mathcal{O}(\epsilon)$.

(b) If $0 < b_* < \frac{N+1}{3}$ then (i) $E^{\gamma^\epsilon}[T_N] = \mathcal{O}(\epsilon^{N-3b_*+1})$; (ii) $E^{\gamma^\epsilon}[T_{B_2}] = \mathcal{O}(\epsilon^{N-3b_*+1})$; (iii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_1}]} = \mathcal{O}(\epsilon^{N-b^*})$ and (iv) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_0]} \le \mathcal{O}(\epsilon^{N-b^*})$.

(c) If $b_* = \frac{N+1}{3}$ then (i) $E^{\gamma^\epsilon}[T_N] = \mathcal{O}(1)$; (ii) $E^{\gamma^\epsilon}[T_{B_2}] = \mathcal{O}(1)$; (iii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_1}]} = \mathcal{O}(\epsilon^{\frac{2N+2}{3}})$ and (iv) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_0]} \le \mathcal{O}(\epsilon^{\frac{2N+2}{3}})$.

(d) Let $\frac{N+1}{3} < b_* < \frac{N+1}{2}$. Then (i) $\frac{1}{E^{\gamma^\epsilon}[T_N]} = \mathcal{O}(\epsilon^{3b_*-N-1})$; (ii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_2}]} = \mathcal{O}(\epsilon^{3b_*-N-1})$; (iii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_1}]} = \mathcal{O}(\epsilon^{N-b^*})$ and (iv) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_0]} \le \mathcal{O}(\epsilon^{N-b^*})$. (v) Moreover, for all $1 \le x \le b^*$, we have $E^{\gamma^\epsilon}[T_{x..N}] = \mathcal{O}(\epsilon^{N+1-2b_*-x})$.

(e) If $b_* = \frac{N+1}{2}$ then (i) $\frac{1}{E^{\gamma^\epsilon}[T_N]} = \mathcal{O}(\epsilon^{\frac{N+1}{2}})$; (ii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_2}]} = \mathcal{O}(\epsilon^{\frac{N+1}{2}})$; (iii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_1}]} = \mathcal{O}(\epsilon^{\frac{N+1}{2}})$; (iv) $\frac{1}{E^{\gamma^\epsilon}[T_0]} = \mathcal{O}(\epsilon^{\frac{N+1}{2}})$ and (v) $\mathcal{O}(\epsilon) \le E^{\gamma^\epsilon}[T_{\{0,N\}}] \le \mathcal{O}(1)$.

(f) Let $\frac{N+1}{2} < b_* < \frac{2N+2}{3}$. Then (i) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_N]} \le \mathcal{O}(\epsilon^{b_*})$; (ii) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_{B_2}]} \le \mathcal{O}(\epsilon^{b_*})$; (iii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_1}]} = \mathcal{O}(\epsilon^{2N-3b_*+2})$ and (iv) $\frac{1}{E^{\gamma^\epsilon}[T_0]} = \mathcal{O}(\epsilon^{2N-3b_*+2})$. (v) Finally, if $b_* \le x \le N-1$ then $E^{\gamma^\epsilon}[T_{0..x}] = \mathcal{O}(\epsilon^{2b_*-2N-1+x})$.

(g) Let $b_* = \frac{2N+2}{3}$. Then (i) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_N]} \le \mathcal{O}(\epsilon^{\frac{2N+2}{3}})$; (ii) $E^{\gamma^\epsilon}[T_{B_2}] = \mathcal{O}(\epsilon^{b_*})$; (iii) $E^{\gamma^\epsilon}[T_{B_1}] = \mathcal{O}(1)$ and (iv) $E^{\gamma^\epsilon}[T_0] = \mathcal{O}(1)$.

(h) Let $b_* > \frac{2N+2}{3}$. Then (i) $\mathcal{O}(\epsilon^N) \le \frac{1}{E^{\gamma^\epsilon}[T_N]} \le \mathcal{O}(\epsilon^{b_*})$; (ii) $\frac{1}{E^{\gamma^\epsilon}[T_{B_2}]} = \mathcal{O}(\epsilon^{b_*})$; (iii) $E^{\gamma^\epsilon}[T_{B_1}] = \mathcal{O}(\epsilon^{3b_*-2N-2})$ and (iv) $E^{\gamma^\epsilon}[T_0] = \mathcal{O}(\epsilon^{3b_*-2N-2})$.

Proof. We first prove (a) and (b)–(d)(i). Note that, without loss of generality, we may assume that $(N-1)..N \subseteq B_2$. Otherwise, $B_2 = \{N\}$ and we are considering the cases (b)–(d)(ii).

Consider the transition matrix $(K^\epsilon)^{\gamma^\epsilon,\{N\}}$, cf. (11), i.e.,

$$(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x,\cdot\,) = \begin{cases} K^\epsilon(x,\cdot\,) & , \quad x \in E \setminus \{N\}, \\ \gamma^\epsilon(\cdot) & , \quad x = N. \end{cases} \tag{30}$$

Since all the elements in $(K^\epsilon)^{\gamma^\epsilon,\{N\}}$ are positive (γ^ϵ is the stationary distribution of the irreducible Markov kernel K^ϵ), this transition matrix has a unique fixed probability vector β^ϵ. By 3.6, we obtain

$$E^{\gamma^\epsilon}[T_N] = \frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)}. \tag{31}$$

We now prove that, if $0 \le b_* < \frac{N+1}{2}$ then, for all $x \in b_*..(N-1)$, we have

$$\mathcal{O}(\epsilon^{N-x}) \le \frac{\beta^\epsilon(x)}{\beta^\epsilon((x+1)..N)} \le \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x\}}), \tag{32}$$

and

$$\mathcal{O}(\epsilon^{N-y}) \le \beta^\epsilon_{x..N}(y) \le \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-y\}}), \qquad y \in x..N. \tag{33}$$

For this sake, we use (backward) mathematical induction. By (5) and (7) applied to the case $B = \{N\}$ and $C = \{N-1\}$, we obtain

$$\frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)}{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N)} \le \frac{\beta^\epsilon(N)}{\beta^\epsilon(N-1)} \le \frac{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N-1)}. \tag{34}$$

By 5.4(ii), we have $\gamma^\epsilon(0..(N-1)) = \mathcal{O}(\epsilon)$ and $\gamma^\epsilon(N-1) = \mathcal{O}(\epsilon)$ and by (20) and (21) together with $b(N-1) = N$ (recall that $N-1 \in B_2$), we have

$$(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N) = P[S^\epsilon(b(N-1)) = N] = P[S^\epsilon(N) = N] = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)} = \mathcal{O}(\epsilon)$$

and

$$\begin{aligned} & 1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N-1) \\ = \;& (K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N) + (K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,0..(N-2)) \\ = \;& P[S^\epsilon(N) = N-1] + P[S^\epsilon(N) \le N-2] \\ = \;& \mathcal{O}(1) + \mathcal{O}(\epsilon^2) = \mathcal{O}(1). \end{aligned}$$

These assertions together with the Definition (30) imply

$$\frac{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,0..(N-1))}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \frac{\gamma^\epsilon(0..(N-1))}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \mathcal{O}(\epsilon),$$

and

$$\begin{aligned} & \frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N-1)}{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N-1)} \\ = \;& \frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N) + (K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,0..(N-2))} \\ = \;& \frac{\gamma^\epsilon(N-1)}{\mathcal{O}(1)+\mathcal{O}(\epsilon^2)} = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)+\mathcal{O}(\epsilon^2)} = \mathcal{O}(\epsilon). \end{aligned}$$

These two facts together with (34) imply $\frac{\beta^\epsilon(N-1)}{\beta^\epsilon(N)} = \mathcal{O}(\epsilon)$. Following the same reasoning as the one used in the proof of Proposition 5.2, we obtain $\beta^\epsilon_{(N-1)..N}(N-1) = \mathcal{O}(\epsilon)$ and $\beta^\epsilon_{(N-1)..N}(N) = \mathcal{O}(1)$. Hence, we have proved (32) and (33), for $x = N-1$. Suppose now that both assertions are already proved, for all $x \geq b_* + 1$. By (3) and (6) applied to the case $B = x..N$ and $C = x-1$, we obtain

$$\frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x..N)}{1-\sum_{y=x}^{N}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,x..N)} \leq \frac{\beta^\epsilon(x..N)}{\beta^\epsilon(x-1)} \leq \frac{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x-1)}{\sum_{y=x}^{N}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,x-1)}. \tag{35}$$

Now, by 5.4(ii) together with the definition of the matrix $(K^\epsilon)^{\gamma^\epsilon,\{N\}}$, cf. (30), we have next $(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,x-1) = \gamma^\epsilon(x-1) = \mathcal{O}(\epsilon^{N-x+1})$. Moreover, since $x \geq b_*$, we have $b(x-1) \geq x \geq x-2$ and, therefore, by (21) and (22) we obtain $P[S^\epsilon(b(x-1)) \leq x-2] = \mathcal{O}(\epsilon^{b(x-1)-x+2})$ and $P[S^\epsilon(b(x-1)) \geq x] = \mathcal{O}(1)$, respectively. Furthermore, since for all $y \in x..(N-1)$, we have $b(y) \geq x+1 \geq x-1$, it follows from (20) that $P[S^\epsilon(b(y)) = x] = \mathcal{O}(\epsilon^{b(y)-x})$. These assertions together with the induction hypothesis (left hand side of (33)) and the definition of the matrix $(K^\epsilon)^{\gamma^\epsilon,\{N\}}$ imply

$$\begin{aligned}
& \frac{\sum_{y=x}^{N}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,x-1)}{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x-1)} \\
= & \frac{\sum_{y=x}^{N-1}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,x-1)+\beta^\epsilon_{x..N}(N)\cdot\gamma^\epsilon(x-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,0..(x-2))+(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x..N)} \\
\geq & \frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{N-y})P[S^\epsilon(b(y))=x-1]+\mathcal{O}(1)\cdot\mathcal{O}(\epsilon^{N-x+1})}{P[S^\epsilon(b(x-1))\leq x-2]+P[S^\epsilon(b(x-1))\geq x]} \\
= & \frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{N-y})P[S^\epsilon(b(y))=x-1]+\mathcal{O}(1)\cdot\mathcal{O}(\epsilon^{N-x+1})}{\mathcal{O}(\epsilon^{b(x-1)-x+1})+\mathcal{O}(1)} \\
= & \frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{N-y})\mathcal{O}(\epsilon^{b(y)-x+1})+\mathcal{O}(\epsilon^{N-x+1})+\beta^\epsilon_{x..N}(N)\gamma^\epsilon(x-1)}{\mathcal{O}(1)} \\
= & \frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{N-y+b(y)-x+1})+\mathcal{O}(\epsilon^{N-x+1})}{\mathcal{O}(1)} = \frac{\mathcal{O}(\epsilon^{N-x+1})}{\mathcal{O}(1)} = \mathcal{O}(\epsilon^{N-x+1}),
\end{aligned} \tag{36}$$

where the last but one assertion follows from the fact that $b(y)-y \geq 1$ for all $y \in x..(N-1)$ implies $N-y+b(y)-x+1 > N-x+1$ and, therefore, $\min\{\min\{N-y+b(y)-x+1 \;/\; y \in x..(N-1)\}, N-x+1\} = N-x+1$. Using this time the right hand side of (33) and following an analogous argument as above, we obtain

$$
\begin{aligned}
& \frac{1-\sum_{y=x}^{N} \beta^{\epsilon}_{x..N}(y)(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(y,x..N)}{(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(x-1,x..N)} \\
\leq\ & \frac{\sum_{y=x}^{N-1} \beta^{\epsilon}_{x..N}(y)(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(y,0..(x-1)) + \beta^{\epsilon}_{x..N}(N)(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(N,0..(x-1))}{(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(x-1,x..N)} \\
\leq\ & \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{\min\{N-y,N-2b_*+1\}})P[S^{\epsilon}(b(y)) \leq x-1] + \mathcal{O}(1)\cdot\gamma^{\epsilon}(0..(x-1))}{P[S^{\epsilon}(b(x-1)) \geq x]} \\
=\ & \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{\min\{N-y,N-2b_*+1\}})\mathcal{O}(\epsilon^{b(y)-x+1}) + \gamma^{\epsilon}(0..(x-1))}{\mathcal{O}(1)} \\
=\ & \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{\min\{N-y,N-2b_*+1\}+b(y)-x+1}) + \gamma^{\epsilon}(0..(x-1))}{\mathcal{O}(1)}.
\end{aligned}
$$

Now, by 5.4(ii), we have

$$
\begin{aligned}
& \gamma^{\epsilon}(0..(x-1)) = \gamma^{\epsilon}(B_1) + \gamma^{\epsilon}(b_*..(x-1)) \\
=\ & \mathcal{O}(\epsilon^{N-2b_*+1}) + \mathcal{O}(\epsilon^{N-x+1}) = \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}}).
\end{aligned}
$$

Moreover, since for all $y \in x..(N-1)$ we have $b(y)-y \geq 1$ and $b(y)-x+1 \geq 2$, hence

$$
\begin{aligned}
& \min\{N-y, N-2b_*+1\} + b(y) - x + 1 \\
=\ & \min\{N-y+b(y)-x+1, N-2b_*+1+b(y)-x+1\} \\
\geq\ & \min\{N-x+1+b(y)-y, N-2b_*+1\} \\
\geq\ & \min\{N-x+1, N-2b_*+1\},
\end{aligned}
$$

for all $y \in x..(N-1)$. The last three assertions imply

$$
\begin{aligned}
\frac{1-\sum_{y=x}^{N} \beta^{\epsilon}_{x..N}(y)(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(y,x..N)}{(K^{\epsilon})^{\gamma^{\epsilon},\{N\}}(x-1,x..N)} & \leq \frac{\mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}})}{\mathcal{O}(1)} \\
& = \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}}). \qquad (37)
\end{aligned}
$$

Thus, (35), (36), and (37) imply Claim (32). Simple algebraic manipulations applied to this claim infer in turn

$$
\mathcal{O}(\epsilon^{N-x+1}) \leq \beta^{\epsilon}_{(x-1)..N}(x-1) \leq \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}})
$$

and

$$
\beta^{\epsilon}_{(x-1)..N}(x..N) = \mathcal{O}(1). \qquad (38)
$$

Now, we may write the probability vector $\beta^{\epsilon}_{(x-1)..N}$ as

$$
\beta^{\epsilon}_{(x-1)..N} = \beta^{\epsilon}_{(x-1)..N}(x-1)\delta_{x-1} + \beta^{\epsilon}_{(x-1)..N}(x..N)\beta^{\epsilon}_{x..N}. \qquad (39)
$$

Hence, for all $y \in x..N$, we have $\beta^{\epsilon}_{(x-1)..N}(y) = \beta^{\epsilon}_{(x-1)..N}(x..N)\beta^{\epsilon}_{x..N}(y)$. Using this equality together with (38) and the induction hypothesis (33), we obtain

$$
\beta^{\epsilon}_{(x-1)..N}(y) \leq \mathcal{O}(1)\cdot\mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}}) = \mathcal{O}(\epsilon^{\min\{N-2b_*+1,N-x+1\}})
$$

and

$$\beta^\epsilon_{(x-1)..N}(y) \geq \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^{N-y}) = \mathcal{O}(\epsilon^{N-y}).$$

As a corollary of this assertion, we have that, if $b_* < \frac{N+1}{2}$ then

$$\mathcal{O}(\epsilon^{N-x}) \leq \beta^\epsilon_{B_2}(x) \leq \mathcal{O}(\epsilon^{\min\{N-2b_*+1, N-x\}}), \qquad x \in B_2. \tag{40}$$

Now, if $b_* = 0$ then (40) implies $\beta^\epsilon(x) = \mathcal{O}(\epsilon^{N-x})$ for all $x \in E$. Therefore, we obtain $\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)} = \mathcal{O}(\epsilon)$. Claim (a) follows now from (31). If, otherwise, $b_* > 0$ then by Proposition 5.2(b) applied to the transition matrix $Q^\epsilon = (K^\epsilon)^{\gamma^\epsilon,\{N\}}$ and $m = b^*$, we obtain

$$\beta^\epsilon_{B_1}(x) = \mathcal{O}(\epsilon^x), \qquad x \in B_1. \tag{41}$$

By (3), applied to the case $B = B_1$, $C = B_2$, and $Q = (K^\epsilon)^{\gamma^\epsilon,\{N\}}$, we have

$$\frac{\beta^\epsilon(B_1)}{\beta^\epsilon(B_2)} = \frac{\sum_{y\in B_2} \beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y, B_1)}{\sum_{x\in B_1} \beta^\epsilon_{B_1}(x)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x, B_2)}. \tag{42}$$

Moreover, by (41) together with (30) and the fact $\min_{x\in B_1}(x - b(x)) = 0$, we obtain

$$\begin{aligned}
&\sum_{x\in B_1} \beta^\epsilon_{B_1}(x)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x, B_2) \\
= &\sum_{x\in B_1} \mathcal{O}(\epsilon^x) P[S^\epsilon(b(x)) \geq b_*] = \sum_{x\in B_1} \mathcal{O}(\epsilon^x) \cdot \mathcal{O}(\epsilon^{b_*-b(x)}) \\
= &\sum_{x\in B_1} \mathcal{O}(\epsilon^{x+b_*-b(x)}) = \mathcal{O}(\epsilon^{b_*}).
\end{aligned} \tag{43}$$

Now, by the left hand side of (40) together with the assertions $(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N, B_1) = \gamma^\epsilon(B_1) = \mathcal{O}(\epsilon^{N-2b_*+1})$, cf. 5.4(ii), $P[S^\epsilon(b(y)) \in B_1] = \mathcal{O}(\epsilon^{b(y)-b^*})$ for all $y \in B_2$ and $\min_{x\in b_*..(N-1)}(N - x + b(x) - b^*) = N + 1 - b^*$, we obtain

$$\begin{aligned}
&\sum_{y\in B_2} \beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y, B_1) \\
\geq &\sum_{y=b_*}^{N-1} \mathcal{O}(\epsilon^{N-y}) P[S^\epsilon(b(y)) \in B_1] + \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^{N-2b_*+1}) \\
= &\sum_{y=b_*}^{N-1} \mathcal{O}(\epsilon^{N-y+b(y)-b^*}) + \mathcal{O}(\epsilon^{N-2b_*+1}) \\
= &\ \mathcal{O}(\epsilon^{N-b^*+1}) + \mathcal{O}(\epsilon^{N-2b_*+1}) = \mathcal{O}(\epsilon^{N-2b_*+1}).
\end{aligned} \tag{44}$$

Similarly, by the right hand side of (40), we obtain

$$\begin{aligned}
&\sum_{y\in B_2} \beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y, B_1) \\
\leq &\sum_{y=b_*}^{N-1} \mathcal{O}(\epsilon^{\min\{N-2b_*+1, N-y\}}) \cdot \mathcal{O}(\epsilon^{b(y)-b^*}) + \mathcal{O}(1) \cdot \mathcal{O}(\epsilon^{N-2b_*+1}) \\
= &\ \mathcal{O}(\epsilon^{N-2b_*+1}),
\end{aligned} \tag{45}$$

where the last equality follows from the fact that, for all $y \in b_*..(N-1)$, we have

$$\min\{N - 2b_* + 1, N - y\} + b(y) - b^* \geq N - 2b_* + 1.$$

By (42)–(45), we obtain

$$\frac{\beta^\epsilon(B_1)}{\beta^\epsilon(B_2)} = \frac{\mathcal{O}(\epsilon^{N-2b_*+1})}{\mathcal{O}(\epsilon^{b_*})} = \mathcal{O}(\epsilon^{N-3b_*+1}). \tag{46}$$

Let us now consider the case $b_* \leq \frac{N+1}{3}$ or, equivalently, $N - 3b_* + 1 > 0$. The assertion (46) implies $\beta^\epsilon(B_1) = \mathcal{O}(\epsilon^{N-3b_*+1})$ and $\beta^\epsilon(B_2) = \mathcal{O}(1)$. This fact together with (40) and (41) implies

$$\beta^\epsilon(x) = \beta^\epsilon(B_1) \cdot \beta^\epsilon_{B_1}(x) = \mathcal{O}(\epsilon^{N-3b_*+1}) \cdot \mathcal{O}(\epsilon^x) = \mathcal{O}(\epsilon^{N-3b_*+1+x}), \qquad x \in B_1,$$

and

$$\mathcal{O}(\epsilon^{N-x}) \leq \beta^\epsilon(x) \leq \mathcal{O}(\epsilon^{\min\{N-2b_*+1, N-x\}}), \qquad x \in B_2. \tag{47}$$

Hence,

$$\begin{aligned} \frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} &= \frac{\beta^\epsilon(B_1) + \beta^\epsilon(b_*..(N-1))}{\beta^\epsilon(N)} \\ &\leq \frac{\mathcal{O}(\epsilon^{N-3b_*+1}) + \mathcal{O}(\epsilon^{\min\{N-2b_*+1),1\}})}{\mathcal{O}(1)}. \end{aligned} \tag{48}$$

Now, if $b_* = \frac{N+1}{3}$ then (48) implies $\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} \leq \mathcal{O}(1)$. If, otherwise, $b_* < \frac{N+1}{3}$ then (48) implies $\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} \leq \mathcal{O}(\epsilon)$. A similar procedure shows that the left hand side of (47) implies $\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} \geq \mathcal{O}(\epsilon)$, if $b_* < \frac{N+1}{3}$ and $\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} \geq \mathcal{O}(1)$, if $b_* = \frac{N+1}{3}$. Therefore,

$$\frac{\beta^\epsilon(0..(N-1))}{\beta^\epsilon(N)} = \begin{cases} \mathcal{O}(\epsilon), & b_* < \frac{N+1}{3}, \\ \mathcal{O}(1), & b_* = \frac{N+1}{3}. \end{cases}$$

Claims (b) and (c)(i) follow now from (31).

We prove next (d)(i). Here, we have $3b_* - N - 1 > 0$. From (46), it follows $\beta^\epsilon(B_2) = \mathcal{O}(\epsilon^{3b_*-N-1})$ and $\beta^\epsilon(B_1) = \mathcal{O}(1)$. Hence, by (40), we obtain

$$\mathcal{O}(\epsilon^{3b_*-x-1}) \leq \beta^\epsilon(x) \leq \mathcal{O}(\epsilon^{\min\{N-2b_*+1, N-x\}+3b_*-N-1}),$$

for all $x \in B_2$. Hence, $\beta^\epsilon(N) = \mathcal{O}(\epsilon^{3b_*-N-1})$ and $\beta^\epsilon(x) \geq \mathcal{O}(\epsilon^{r_x})$ for an $r_x \geq 2$ and for all $x \in B_2 \setminus \{N\}$. Resuming these last assertions, we obtain

$$\frac{\beta^\epsilon(N)}{\beta^\epsilon(0..(N-1))} = \frac{\beta^\epsilon(N)}{\beta^\epsilon(B_1) + \beta^\epsilon(B_2 \setminus \{N\})} = \frac{\mathcal{O}(\epsilon^{3b_*-N-1})}{\mathcal{O}(1)} = \mathcal{O}(\epsilon^{3b_*-N-1}).$$

The claim (d)(i) follows now from (31). In order to prove (e)(i), we assume, as before, $(N-1)..N \subseteq B_2$. Our first claim is that, if $b_* = \frac{N+1}{2}$ then we have

$$\mathcal{O}(\epsilon^{N-x}) \leq \frac{\beta^\epsilon(x)}{\beta^\epsilon((x+1)..N)} \leq \mathcal{O}(1) \tag{49}$$

for all $x \in b_*..(N-1)$ and

$$\mathcal{O}(\epsilon^{N-y}) \leq \beta^\epsilon_{x..N}(y) \leq \mathcal{O}(1), \qquad y \in x..N. \tag{50}$$

The proof of these two assertions is similar to the proof of (32) and (33). Therefore, we confine ourselves to sketch it. By 5.4(iii), we obtain

$$\frac{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,0..(N-1))}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \frac{\gamma^\epsilon(0..(N-1))}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} = \mathcal{O}(1)$$

and

$$\begin{aligned}
& \frac{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,N-1)}{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N-1)} \\
= \; & \frac{\gamma^\epsilon(N-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,0..(N-2))+(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N-1,N)} \\
= \; & \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)} = \mathcal{O}(\epsilon).
\end{aligned}$$

These last two assertions together with (34) imply $\mathcal{O}(\epsilon) \leq \frac{\beta^\epsilon(N-1)}{\beta^\epsilon(N)} \leq \mathcal{O}(1)$ and, therefore, $\mathcal{O}(\epsilon) \leq \beta^\epsilon_{(N-1)..N}(N) \leq \mathcal{O}(1)$ and $\beta^\epsilon_{(N-1)..N}(N-1) = \mathcal{O}(1)$. This last assertion together with (34) implies $\mathcal{O}(\epsilon) \leq \frac{\beta^\epsilon(N)}{\beta^\epsilon(N-1)} \leq \mathcal{O}(1)$ and, therefore, $\beta^\epsilon_{(N-1)..N}(N) = \mathcal{O}(1)$ and $\mathcal{O}(\epsilon) \leq \beta^\epsilon_{(N-1)..N}(N-1) \leq \mathcal{O}(1)$. This concludes the proof of (49) and (50) for $x = N-1$. Suppose now that our assertions are already proved for all $x \geq b_* + 1$. Using the induction hypothesis, we obtain

$$\begin{aligned}
& \frac{\sum_{y=x}^{N-1} \beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,0..(x-1)) + \beta^\epsilon_{x..N}(N)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,0..(x-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x..N)} \\
\leq \; & \frac{\sum_{y=x}^{N-1} \mathcal{O}(1)\cdot P[S^\epsilon(b(y)) \leq x-1] + \mathcal{O}(1)\cdot\gamma^\epsilon(0..(x-1))}{P[S^\epsilon(b(x-1)) \leq x-2] + P[S^\epsilon(b(x-1)) \geq x]} \\
= \; & \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{b(y)-x+1}) + \gamma^\epsilon(0..(x-1))}{\mathcal{O}(1)} \\
= \; & \frac{\sum_{y=x}^{N-1} \mathcal{O}(\epsilon^{b(y)-x+1}) + \mathcal{O}(1)}{\mathcal{O}(1)} \\
= \; & \mathcal{O}(1),
\end{aligned}$$

where the last assertion follows from the fact that, for all $y \in x..(N-1)$, we have $b(y) - x \geq 1$. Following a similar procedure as the one used to obtain (36) but using this time 5.4(iii), we have

$$
\begin{aligned}
&\frac{\sum_{y=x}^{N-1}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,0..(x-1))+\beta^\epsilon_{x..N}(N)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(N,0..(x-1)}{(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x..N)}\\
\leq\;&\frac{\sum_{y=x}^{N-1}\mathcal{O}(1)\cdot P[S^\epsilon(b(y))\leq x-1]+\mathcal{O}(1)\cdot\gamma^\epsilon(0..(x-1))}{P[S^\epsilon(b(x-1))\leq x-2]+P[S^\epsilon(b(x-1))\geq x]}\\
=\;&\frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{b(y)-x+1})+\gamma^\epsilon(0..(x-1))}{\mathcal{O}(1)}\\
=\;&\frac{\sum_{y=x}^{N-1}\mathcal{O}(\epsilon^{b(y)-x+1})+\mathcal{O}(1)}{\mathcal{O}(1)}\\
=\;&\mathcal{O}(1),
\end{aligned}
$$

and

$$
\frac{\sum_{y=x}^{N}\beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,0..(x-1))}{1-(K^\epsilon)^{\gamma^\epsilon,\{N\}}(x-1,x-1)}\geq\mathcal{O}(\epsilon^{N-x+1}).
$$

Hence, by (35) together with the last two inequalities, we obtain (49) for $x-1$. This assertion in turn implies $\mathcal{O}(\epsilon^{N-x+1})\leq\beta^\epsilon_{(x-1)..N}(x-1)\leq\mathcal{O}(1)$ and $\beta^\epsilon_{(x-1)..N}(x..N)=\mathcal{O}(1)$. Claim (50) follows now from these last assertions together with (39). As a corollary from (50), we obtain that, if $b_*=\frac{N+1}{2}$ then

$$
\mathcal{O}(\epsilon^{N-x})\leq\beta^\epsilon_{B_2}(x)\leq\mathcal{O}(1),\qquad x\in B_2. \tag{51}
$$

Therefore, from the assertion above together with $\gamma^\epsilon(B_1)=\mathcal{O}(1)$, cf. 5.4(iii), we may infer

$$
\begin{aligned}
&\sum_{y\in B_2}\beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,B_1)\\
=\;&\sum_{y=b_*}^{N-1}\beta^\epsilon_{B_2}(y)P[S^\epsilon(b(y))\in B_1]+\beta^\epsilon_{B_2}(N)\gamma^\epsilon(B_1)\\
\leq\;&\sum_{y=b_*}^{N-1}\mathcal{O}(1)\cdot\mathcal{O}(\epsilon^{b(y)-b^*})+\mathcal{O}(1)\cdot\mathcal{O}(1)\\
=\;&\mathcal{O}(1),
\end{aligned}
$$

where the last equality follows from the fact that $b(y)-b^*\geq 1$ for all $y\in B_2\setminus\{N\}$. Moreover,

$$
\begin{aligned}
&\sum_{y\in B_2}\beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{N\}}(y,B_1)\\
\geq\;&\sum_{y=b_*}^{N-1}\mathcal{O}(\epsilon^{N-y})\cdot\mathcal{O}(\epsilon^{b(y)-b^*})+\mathcal{O}(1)\\
=\;&\sum_{y=b_*}^{N-1}\mathcal{O}(\epsilon^{N-y+b(y)-b^*})+\mathcal{O}(1)\\
=\;&\mathcal{O}(1),
\end{aligned}
$$

The last two assertions together with (42) and (43) imply $\frac{\beta^\epsilon(B_2)}{\beta^\epsilon(B_1)} = \mathcal{O}(\epsilon^{b_*})$. Hence, $\beta^\epsilon(B_1) = \mathcal{O}(1)$ and $\beta^\epsilon(B_2) = \mathcal{O}(\epsilon^{b_*})$. Resuming, we obtain $\mathcal{O}(\epsilon^{N-x+b_*}) \leq \beta^\epsilon(x) \leq \mathcal{O}(\epsilon^{b_*})$, $x \in B_2$, and $\beta^\epsilon(x) = \mathcal{O}(\epsilon^x)$, $x \in B_1$. Therefore, $\frac{\beta^\epsilon(N)}{\beta^\epsilon(0..(N-1))} = \mathcal{O}(\epsilon^{b_*})$. The claim follows now from (31).

For the proof of (b), (c), and (d)(ii), consider the transition matrix $(K^\epsilon)^{\gamma^\epsilon, B_2}$, cf. (11), i.e.,

$$(K^\epsilon)^{\gamma^\epsilon, B_2}(x, \cdot\,) = \begin{cases} K^\epsilon(x, \cdot\,), & x \in B_1, \\ \gamma^\epsilon(\cdot\,) \quad , & x \in B_2. \end{cases}$$

Since all the elements in $(K^\epsilon)^{\gamma^\epsilon, B_2}$ are positive this transition matrix has a unique fixed probability vector which we denote again by β^ϵ. By Proposition 3.6 together with (3) applied to the case $Q = (K^\epsilon)^{\gamma^\epsilon, B_2}$, $B = B_1$ and $C = B_2$, we obtain

$$\begin{aligned} E^{\gamma^\epsilon}[T_{B_2}] &= \frac{\beta^\epsilon(B_1)}{\beta^\epsilon(B_2)} \\ &= \frac{\sum_{y \in B_2} \beta^\epsilon_{B_2}(y)(K^\epsilon)^{\gamma^\epsilon, B_2}(y, B_1)}{\sum_{x \in B_1} \beta^\epsilon_{B_1}(x)(K^\epsilon)^{\gamma^\epsilon, B_2}(x, B_2)} \\ &= \frac{\gamma^\epsilon(B_1)}{\sum_{x \in B_1} \beta^\epsilon_{B_1}(x) P[S^\epsilon(b(x)) \in B_2]}. \end{aligned} \tag{52}$$

Here, we also have $\beta^\epsilon_{B_1}(x) = \mathcal{O}(\epsilon^x)$ for all $x \in B_1$. Moreover, by 5.4(ii), we have $\gamma^\epsilon(B_1) = \mathcal{O}(\epsilon^{N-2b_*+1})$ and by (20), we have $P[S^\epsilon(b(x)) \in B_2] = \mathcal{O}(\epsilon^{b_*-b(x)})$. These facts together with (52) imply

$$\begin{aligned} E^{\gamma^\epsilon}[T_{B_2}] &= \frac{\mathcal{O}(\epsilon^{N-2b_*+1})}{\sum_{x \in B_1} \mathcal{O}(\epsilon^x) \cdot \mathcal{O}(\epsilon^{b_*-b(x)})} \\ &= \frac{\mathcal{O}(\epsilon^{N-2b_*+1})}{\sum_{x \in B_1} \mathcal{O}(\epsilon^{x-b(x)+b_*})} \\ &= \frac{\mathcal{O}(\epsilon^{N-2b_*+1})}{\mathcal{O}(\epsilon^{b_*})} \\ &= \mathcal{O}(\epsilon^{N-3b_*+1}). \end{aligned}$$

This is the claim.

The proof of (e)(ii) is similar as the one given above. However, in this case, we have $\gamma^\epsilon(B_1) = \mathcal{O}(1)$, cf. 5.4(iii).

In order to prove (b), (c), (d) and (e)(iii), consider the transition matrix

$$(K^\epsilon)^{\gamma^\epsilon, B_1}(x, \cdot\,) = \begin{cases} \gamma^\epsilon(\,\cdot\,) \quad , & x \in B_1, \\ K^\epsilon(\,x, \cdot\,) \quad , & x \in B_2. \end{cases} \tag{53}$$

This transition matrix is irreducible and, therefore, has a unique stationary distribution β^ϵ. By Proposition 3.6 together with (3) applied to the case $Q = (K^\epsilon)^{\gamma^\epsilon, B_1}$, $B = B_2$, and $C = B_1$

and (53), we have

$$
\begin{aligned}
E^{\gamma^\epsilon}[T_{B_1}] &= \frac{\beta^\epsilon(B_2)}{\beta^\epsilon(B_1)} \\
&= \frac{\sum_{x\in B_1}\beta^\epsilon_{B_1}(x)(K^\epsilon)^{\gamma^\epsilon,B_1}(x,B_2)}{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)(K^\epsilon)^{\gamma^\epsilon,B_1}(x,B_1)} \\
&= \frac{\gamma^\epsilon(B_2)}{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)P[S^\epsilon(b(x))\in B_1]} \\
&= \frac{\gamma^\epsilon(B_2)}{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)\mathcal{O}(\epsilon^{b(x)-b^*})}, \qquad (54)
\end{aligned}
$$

where the last equality follows from (20) together with the fact $b(x)\geq b_*+1\geq b^*$ for all $x\in B_2$. Now, if $b_*\leq\frac{N+1}{2}$ then from 5.4(ii)–(iii) it follows $\gamma^\epsilon(B_2)=\mathcal{O}(1)$. Moreover, by Proposition 5.2(a) applied to $Q^\epsilon=(K^\epsilon)^{\gamma^\epsilon,B_1}$, we obtain $\beta^\epsilon_{B_2}(x)=\mathcal{O}(\epsilon^{N-x})$ for all $x\in B_2$. These assertions together with (54) imply

$$
\begin{aligned}
E^{\gamma^\epsilon}[T_{B_1}] &= \frac{\mathcal{O}(1)}{\sum_{x\in B_2}\mathcal{O}(\epsilon^{N-x})\cdot\mathcal{O}(\epsilon^{b(x)-b^*})} \\
&= \frac{\mathcal{O}(1)}{\sum_{x\in B_2}\mathcal{O}(\epsilon^{N-x+b(x)-b^*})} \\
&= \frac{1}{\mathcal{O}(\epsilon^{N-b^*})}.
\end{aligned}
$$

This is the claim.
We now prove (b)–(e)(iv). For this sake, consider the transition kernel

$$
(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,\cdot\,)=\begin{cases}\gamma^\epsilon(\,\cdot\,) & , \quad y=0,\\ K^\epsilon(\,y,\cdot\,) & , \quad y\in 1..N,\end{cases} \qquad (55)
$$

and, as before, let us denote by β^ϵ its unique stationary distribution. By Proposition 3.6 together with (3) applied to the case $Q=(K^\epsilon)^{\gamma^\epsilon,\{0\}}$, $B=1..N$ and $C=\{0\}$, we first have

$$
E^{\gamma^\epsilon}[T_0]=\frac{\beta^\epsilon(1..N)}{\beta^\epsilon(0)}. \qquad (56)
$$

By 5.2(a) applied to the transition matrix $Q^\epsilon=(K^\epsilon)^{\gamma^\epsilon,\{0\}}$, we obtain

$$
\beta^\epsilon_{B_2}(x)=\mathcal{O}(\epsilon^{N-x}), \qquad x\in B_2. \qquad (57)
$$

Moreover, by (3) and (6) applied to the case $B=B_2$ and $C=\{0\}$, we obtain

$$
\frac{(K^\epsilon)^{\gamma^\epsilon,\{0\}}(0,B_2)}{1-\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,B_2)}\leq\frac{\beta^\epsilon(B_2)}{\beta^\epsilon(0)}\leq\frac{1-(K^\epsilon)^{\gamma^\epsilon,\{0\}}(0,0)}{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,0)}. \qquad (58)
$$

From (55), (57) and (20), we obtain

$$
\begin{aligned}
\frac{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,0)}{1-(K^\epsilon)^{\gamma^\epsilon,\{0\}}(0,0)} &= \frac{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)P[S^\epsilon(b(x))=0]}{\gamma^\epsilon(1..N)} \\
= \frac{\sum_{x\in B_2}\mathcal{O}(\epsilon^{N-x})\cdot\mathcal{O}(\epsilon^{b(x)})}{\gamma^\epsilon(1..N)} &= \frac{\mathcal{O}(\epsilon^N)}{\gamma^\epsilon(1..N)}. \qquad (59)
\end{aligned}
$$

Following similar argumentations, we obtain

$$\begin{aligned}
& \frac{\sum_{x\in B_2}\beta^\epsilon_{B_2}(x)P[S^\epsilon(b(x))\le b^*]}{K^{\epsilon,\{0\}}(0,B_2)} = \frac{\sum_{x\in B_2}\mathcal{O}(\epsilon^{N-x})\cdot\mathcal{O}(\epsilon^{b(x)-b^*})}{\gamma^\epsilon(B_2)} \\
= \;& \frac{\sum_{x\in B_2}\mathcal{O}(\epsilon^{N-x+b(x)-b^*})}{\gamma^\epsilon(B_2)} = \frac{\mathcal{O}(\epsilon^{N-b^*})}{\gamma^\epsilon(B_2)}. \qquad (60)
\end{aligned}$$

From (58)–(60), we obtain

$$\frac{\gamma^\epsilon(B_2)}{\mathcal{O}(\epsilon^{N-b^*})} \le \frac{\beta^\epsilon(B_2)}{\beta^\epsilon(0)} \le \frac{\gamma^\epsilon(1..N)}{\mathcal{O}(\epsilon^N)}. \qquad (61)$$

We now prove that, if $0 < b_* \le \frac{N+1}{2}$ then, for all $x \in B_1$, we have

$$\mathcal{O}(\epsilon^{N-x}) \le \frac{\beta^\epsilon(x)}{\beta^\epsilon(\ 0..(x-1)\ \cup\ B_2\ )} \le \mathcal{O}(\epsilon^{N-b^*}), \qquad (62)$$

$$\beta^\epsilon_{0..x\cup B_2}(y) = \mathcal{O}(\epsilon^{N-y}), \qquad y \in B_2, \qquad (63)$$

and

$$\mathcal{O}(\epsilon^{N-y}) \le \beta^\epsilon_{0..x\ \cup\ B_2}(y) \le \mathcal{O}(\epsilon^{N-b^*}), \qquad y \in 0..x. \qquad (64)$$

For this sake, we use again mathematical induction. By 5.4(ii) together with (61), we obtain $\mathcal{O}(\epsilon^N) \le \frac{\beta^\epsilon(0)}{\beta^\epsilon(B_2)} \le \mathcal{O}(\epsilon^{N-b^*})$. Therefore, $\mathcal{O}(\epsilon^N) \le \beta^\epsilon_{B_2\cup\{0\}}(0) \le \mathcal{O}(\epsilon^{N-b^*})$ and $\beta^\epsilon_{B_2\cup\{0\}}(B_2) = \mathcal{O}(1)$. Using this last assertion together with (57) and the equality

$$\beta^\epsilon_{0\cup B_2} = \beta^\epsilon_{0\cup B_2}(0)\delta_0 + \beta^\epsilon_{0\cup B_2}(B_2)\beta^\epsilon_{B_2}.$$

we obtain $\beta^\epsilon_{B_2\cup\{0\}}(y) = \mathcal{O}(\epsilon^{N-y})$, $y \in B_2$. Hence, we have proved (62)–(64) for $x = 0$. Let now $1 \le x \le b^*$. By (3) and (6), applied to the case $B = 0..(x-1) \cup B_2$ and $C = \{x\}$, we obtain

$$\begin{aligned}
& \frac{(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x, 0..(x-1)\cup B_2)}{1 - \sum_{y\in 0..(x-1)\cup B_2}\beta^\epsilon_{0..(x-1)\cup B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y, 0..(x-1)\cup B_2)} \qquad (65) \\
\le \;& \frac{\beta^\epsilon(0..(x-1)\cup B_2)}{\beta^\epsilon(x)} \\
\le \;& \frac{1-(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,x)}{\sum_{y\in 0..(x-1)\cup B_2}\beta^\epsilon_{0..(x-1)\cup B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x)}. \qquad (66)
\end{aligned}$$

For the sake of brevity, we denote by B the set $0..(x-1)\cup B_2$. We now use Definition (55) together with the equality $\gamma^\epsilon(x) = \mathcal{O}(\epsilon^{x+N-2b_*+1})$, cf. 5.4(ii), the assertions $(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x) = P[S^\epsilon(b(y)) = x] = \mathcal{O}(\epsilon^{b(y)-x})$, for all $y \in B_2$, the asymptotics $(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x) = P[S^\epsilon(b(y)) = x] = \mathcal{O}(\epsilon^{x-b(y)})$, for all $y \in 1..(x-1)$, and

$$\begin{aligned}
& (K^\epsilon)^{\gamma^\epsilon,\{0\}}(x, E\setminus\{x\}) \\
= \;& P[S^\epsilon(b(x)) \le x-1] + P[S^\epsilon(b(x)) \ge (x+1)..b^*] + P[S^\epsilon(b(x)) \ge x+1] = \mathcal{O}(1),
\end{aligned}$$

and the induction hypotheses (63) and (64), in order to show that the reciprocal of (66) equals

$$
\begin{aligned}
& \frac{\sum_{y\in B_2}\beta^\epsilon_B(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x)+\sum_{y=1}^{x-1}\beta^\epsilon_B(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x)+\beta^\epsilon_B(0)\gamma^\epsilon(x)}{(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,E\setminus\{x\})} \\
\geq\quad & \frac{\sum_{y\in B_2}\mathcal{O}(\epsilon^{N-y})\cdot\mathcal{O}(\epsilon^{b(y)-x})+\sum_{y=1}^{x-1}\mathcal{O}(\epsilon^{N-y})\cdot\mathcal{O}(\epsilon^{x-b(y)})+\mathcal{O}(\epsilon^{N})\cdot\mathcal{O}(\epsilon^{N-2b_*+1+x})}{\mathcal{O}(1)} \\
=\quad & \frac{\mathcal{O}(\epsilon^{N-x})+\sum_{y=1}^{x-1}\mathcal{O}(\epsilon^{N+x-y-b(y)})+\mathcal{O}(\epsilon^{2N-2b_*+1})}{\mathcal{O}(1)}.
\end{aligned}
\tag{67}
$$

Now, since $b_* \leq \frac{N+1}{2}$, we have $N-y < 2N-2b_*+y+1$ for all $y>0$ and, therefore, $N-x < 2N-2b_*+x+1$. Moreover, for all $y \in 1..(x-1)$, we have $y+b(y) < 2x$ (recall that $x \leq b^*$ implies $b(y) \leq y \leq x-1$) and, therefore, $N-x < N+x-y-b(y)$. These assertions together with imply that (67) $=\mathcal{O}(\epsilon^{N-x})$ and hence, by (66), we have $\frac{\beta^\epsilon(x)}{\beta^\epsilon(0..(x-1)\cup B_2)} \geq \mathcal{O}(\epsilon^{N-x})$, i.e., the left hand side of (62). Now, from (22) (note that $b(x) \leq x-1$), we obtain

$$
(K^\epsilon)^{\gamma^\epsilon,\{0\}}(x,0..(x-1)\cup B_2) = P[S^\epsilon(b(x)) \leq x-1] + P[S^\epsilon(b(x)) \in B_2] = \mathcal{O}(1).
$$

Moreover, from the induction hypotheses (63) and (64) together with the facts $\gamma^\epsilon(x..b^*) = \mathcal{O}(\epsilon^{N-2b_*+1+x})$, cf. 5.4(ii), from the relation

$$
(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x..b^*) = P[S^\epsilon(b(y)) \in x..b^*] = \begin{cases} \mathcal{O}(\epsilon^{x-b(y)}), & y \in 1..b^*, \\ \mathcal{O}(\epsilon^{b(y)-b^*}), & y \in B_2, \end{cases}
$$

cf.(20), and from (57), we obtain

$$
\begin{aligned}
& \frac{1-\sum_{y\in B}\beta^\epsilon_{0..(x-1)\cup B_2}(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,0..(x-1)\cup B_2)}{(K^\epsilon)^{\gamma^\epsilon,0}(x,0..(x-1)\cup B_2)} \\
=\quad & [\ \beta^\epsilon_B(0)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(0,x..b^*)+\sum_{y=1}^{x-1}\beta^\epsilon_B(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x..b^*) \\
& \qquad +\sum_{y\in B_2}\beta^\epsilon_B(y)(K^\epsilon)^{\gamma^\epsilon,\{0\}}(y,x..b^*)\ ]/\mathcal{O}(1) \\
\leq\quad & [\mathcal{O}(\epsilon^{N-b^*})\gamma^\epsilon(x..b^*)+\mathcal{O}(\epsilon^{N-b^*})\sum_{y=1}^{x-1}\mathcal{O}(\epsilon^{x-b(y)}) \\
& \qquad +\sum_{y\in B_2}\mathcal{O}(\epsilon^{N-y})\mathcal{O}(\epsilon^{b(y)-b^*})]/\mathcal{O}(1) \\
=\quad & \frac{\mathcal{O}(\epsilon^{2N-3b_*+2+x})+\mathcal{O}(\epsilon^{N-b^*})\sum_{y=1}^{x-1}\mathcal{O}(\epsilon^{x-b(y)})+\sum_{y\in B_2}\mathcal{O}(\epsilon^{N-y})\mathcal{O}(\epsilon^{b(y)-b^*})}{\mathcal{O}(1)} \\
=\quad & \frac{\mathcal{O}(\epsilon^{2N-3b_*+2+x})+\mathcal{O}(\epsilon^{N-b^*+x-\max_{y\in 1..(x-1)}b(y)})+\mathcal{O}(\epsilon^{N-b^*})}{\mathcal{O}(1)}.
\end{aligned}
\tag{69}
$$

Since $N-2b_*+1 \geq 0$, we have $N-2b_*+1+x > 0$. Therefore, $N-b^* < 2N-3b_*+2+x$. Moreover, by $x \in B_1$ we have $\max_{y\in 1..(x-1)} b(y) \leq x-2$ and, therefore,

$$
N-b^*+x-\max_{y\in 1..(x-1)} b(y) \geq N-b^*+x-(x-2) > N-b^*.
$$

These assertions imply (69) $= \mathcal{O}(\epsilon^{N-x})$; together with (65), we obtain $\mathcal{O}(\epsilon^{N-x}) \le \frac{\beta^\epsilon(x)}{\beta^\epsilon(0..(x-1)\cup B_2)}$. We have thus proved (62) which in turn implies $\mathcal{O}(\epsilon^{N-x}) \le \beta_B^\epsilon(x) \le \mathcal{O}(\epsilon^{N-b^*})$ and $\beta_B^\epsilon(0..(x-1)\cup B_2) = \mathcal{O}(1)$. This last assertion together with the equality

$$\beta^\epsilon_{0..x\cup B_2} = \beta^\epsilon_{0..x\cup B_2}(x)\delta_x + \beta^\epsilon_{0..x\cup B_2}(0..(x-1)\cup B_2)\beta^\epsilon_{0..(x-1)\cup B_2}$$

implies that, for all $y \in 0..(x-1)\cup B_2$, we have $\beta^\epsilon_{0..x\cup B_2}(y) = \mathcal{O}(1)\cdot\beta^\epsilon_{0..(x-1)\cup B_2}(y)$. Hence, the induction hypothesis implies $\mathcal{O}(\epsilon^{N-y}) \le \beta^\epsilon_{0..x\cup B_2}(y) \le \mathcal{O}(\epsilon^{N-b^*})$ if $y \in 0..(x-1)$ and $\beta^\epsilon_{0..x\cup B_2}(y) = \mathcal{O}(\epsilon^{N-y})$ if $y \in B_2$. Thus, we have proved (63) and (64) for x. As a corollary of (63) and (64), we obtain

$$\beta^\epsilon(y) = \mathcal{O}(\epsilon^{N-y}), \quad y \in B_2, \text{ and } \mathcal{O}(\epsilon^{N-y}) \le \beta^\epsilon(y) \le \mathcal{O}(\epsilon^{N-b^*}), \quad y \in B_1.$$

Finally, from (56) together with the claim above, we obtain

$$\begin{aligned}
\frac{1}{E^{\gamma^\epsilon}[T_0]} &= \frac{\beta^\epsilon(0)}{\beta^\epsilon(1..N)} \\
&= \frac{\beta^\epsilon(0)}{\beta^\epsilon(1..b^*)+\beta^\epsilon(B_2)} \ge \frac{\mathcal{O}(\epsilon^N)}{\sum_{x=1}^{b^*}\mathcal{O}(\epsilon^{N-b^*})+\mathcal{O}(1)} \\
&= \frac{\mathcal{O}(\epsilon^N))}{\mathcal{O}(\epsilon^{N-b^*})+\mathcal{O}(1)} = \frac{\mathcal{O}(\epsilon^N))}{\mathcal{O}(1)} = \mathcal{O}(\epsilon^N),
\end{aligned}$$

if $0 < b_* \le \frac{N+1}{2}$ and

$$\begin{aligned}
\frac{1}{E^{\gamma^\epsilon}[T_0]} &= \frac{\beta^\epsilon(0)}{\beta^\epsilon(1..N)} \\
&= \frac{\beta^\epsilon(0)}{\beta^\epsilon(1..b^*)+\beta^\epsilon(B_2)} \le \frac{\mathcal{O}(\epsilon^{N-b^*})}{\sum_{x=1}^{b^*}\mathcal{O}(\epsilon^{N-x})+\mathcal{O}(1)} \\
&= \frac{\mathcal{O}(\epsilon^{N-b^*})}{(\epsilon^{N-b^*})+\mathcal{O}(1)} = \frac{\mathcal{O}(\epsilon^{N-b^*})}{\mathcal{O}(1)} = \mathcal{O}(\epsilon^{N-b^*}).
\end{aligned}$$

We have thus proved (b)–(e)(iv).

We now prove (d)(v). For this sake, let $1 \le x \le b^*$ and consider, as in the cases above, the irreducible transition matrix $(K^\epsilon)^{\gamma^\epsilon,x..N}$, i.e.,

$$(K^\epsilon)^{\gamma^\epsilon,x..N}(y,\cdot\,) = \begin{cases} K^\epsilon(\,y,\cdot\,), & y \in 0..(x-1), \\ \gamma^\epsilon(\,\cdot\,) \quad, & y \in x..N, \end{cases} \tag{70}$$

and, as before, let β^ϵ be its unique stationary distribution. By Proposition 3.6 together with (3) applied to the case $Q = (K^\epsilon)^{\gamma^\epsilon,x..N}$, $B = 0..(x-1)$, and $C = x..N$ and by (70), we have

$$\begin{aligned}
E^{\gamma^\epsilon}[T_{x..N}] = \frac{\beta^\epsilon(0..(x-1))}{\beta^\epsilon(x..N)} &= \frac{\sum_{y=x}^N \beta^\epsilon_{x..N}(y)(K^\epsilon)^{\gamma^\epsilon,x..N}(y,0..(x-1))}{\sum_{y=0}^{x-1}\beta^\epsilon_{0..(x-1)}(y)(K^\epsilon)^{\gamma^\epsilon,x..N}(y,x..N)} \\
&= \frac{\gamma^\epsilon(0..(x-1))}{\sum_{y=0}^{x-1}\beta^\epsilon_{0..(x-1)}(y)P[S^\epsilon(b(y)) \ge x]} \\
&= \frac{\gamma^\epsilon(0..(x-1))}{\sum_{y=0}^{x-1}\beta^\epsilon_{0..(x-1)}(y)\cdot\mathcal{O}(\epsilon^{x-b(y)})},
\end{aligned} \tag{71}$$

where the last equality follows from the left hand side of (21) (note that $0 \leq y \leq x-1 < b^*$ implies $b(y) \leq x$). Now, by 5.2(b) applied to $Q^\epsilon = (K^\epsilon)^{\gamma^\epsilon, x..N}$, we obtain $\beta^\epsilon_{0..(x-1)}(y) = \mathcal{O}(\epsilon^y)$, $y \in 0..(x-1)$. Therefore, from $\min_{y \in 0..(x-1)} y - b(y) = 0$, we obtain

$$\sum_{y=0}^{x-1} \beta^\epsilon_{0..(x-1)}(y) \cdot \mathcal{O}(\epsilon^{x-b(y)}) = \sum_{y=0}^{x-1} \mathcal{O}(\epsilon^y) \cdot \mathcal{O}(\epsilon^{x-b(y)}) = \mathcal{O}(\epsilon^x). \tag{72}$$

Moreover, from 5.4(ii) it follows that $\gamma^\epsilon(0..(x-1)) = \mathcal{O}(\epsilon^{N+1-2b_*})$. This last assertion together with (71) and (72) proves our claim.

We finally prove (e)(v). For this sake, consider the transition matrix $(K^\epsilon)^{\gamma^\epsilon,\{0,N\}}$, i.e.,

$$(K^\epsilon)^{\gamma^\epsilon,\{0,N\}}(y,\cdot\,) = \begin{cases} K^\epsilon(\,y,\cdot\,), & y \in 1..(N-1), \\ \gamma^\epsilon(\,\cdot\,) \quad , & y \in \{0,N\}, \end{cases}$$

and let β^ϵ again be its unique stationary distribution. By 3.6, we have next

$$E^{\gamma^\epsilon}[T_{\{0,N\}}] = \frac{\beta^\epsilon(1..(N-1))}{\beta^\epsilon(\{0,N\})}. \tag{73}$$

Now, the conditions under which (51) was proved are also met in this case. Moreover, by symmetry, we have

$$\mathcal{O}(\epsilon^x) \leq \beta^\epsilon_{B_1}(x) \leq \mathcal{O}(1), \qquad x \in B_1. \tag{74}$$

Further, by (3) applied to the case $Q = (K^\epsilon)^{\gamma^\epsilon,\{0,N\}}$, $B = 1..(N-1)$, and $C = \{0,N\}$ together with 5.4(iii), we have

$$\begin{aligned}
\frac{\beta^\epsilon(B_2)}{\beta^\epsilon(B_1)} &= \frac{\sum_{x\in B_1} \beta^\epsilon_{B_1}(x)(K^\epsilon)^{\gamma^\epsilon,\{0,N\}}(x,B_2)}{\sum_{x\in B_2} \beta^\epsilon_{B_2}(x)(K^\epsilon)^{\gamma^\epsilon,\{0,N\}}(x,B_1)} \\
&= \frac{\sum_{x\in B_1\setminus\{0\}} \mathcal{O}(\epsilon^x)\cdot\mathcal{O}(\epsilon^{b_*-b(x)}) + \mathcal{O}(1)\cdot\gamma^\epsilon(B_2)}{\sum_{x\in B_2\setminus\{0\}} \mathcal{O}(\epsilon^{N-x})\cdot\mathcal{O}(\epsilon^{b(x)-b^*}) + \mathcal{O}(1)\cdot\gamma^\epsilon(B_1)} \\
&= \frac{\sum_{x\in B_1\setminus\{0\}} \mathcal{O}(\epsilon^x)\cdot\mathcal{O}(\epsilon^{b_*-b(x)}) + \mathcal{O}(1)\cdot\gamma^\epsilon(B_2)}{\sum_{x\in B_2\setminus\{0\}} \mathcal{O}(\epsilon^{N-x})\cdot\mathcal{O}(\epsilon^{b(x)-b^*}) + \mathcal{O}(1)\cdot\gamma^\epsilon(B_1)} \\
&= \frac{\mathcal{O}(\epsilon^{b_*}) + \mathcal{O}(1)}{\mathcal{O}(\epsilon^{N-b^*}) + \mathcal{O}(1)} = \mathcal{O}(1).
\end{aligned}$$

Therefore, $\beta^\epsilon(B_2) = \mathcal{O}(1)$ and $\beta^\epsilon(B_1) = \mathcal{O}(1)$. This last assertion together with (51) and (74) implies $\mathcal{O}(\epsilon^x) \leq \beta^\epsilon(x) \leq \mathcal{O}(1)$ for all $x \in B_1$ and $\mathcal{O}(\epsilon^{N-x}) \leq \beta^\epsilon(x) \leq \mathcal{O}(1)$ for all $x \in B_2$. Hence,

$$\mathcal{O}(\epsilon) = \frac{\mathcal{O}(\epsilon)}{\mathcal{O}(1)} \leq \frac{\beta^\epsilon(1..(N-1))}{\beta^\epsilon(\{0,N\})} \leq \frac{\mathcal{O}(1)}{\mathcal{O}(1)} = \mathcal{O}(1).$$

The claim now follows from (73).
Claims (f)–(h) follow from the paragraph at the end of Subsection 4.1 together with Parts (f)–(h). $\square$

Consider the comment before this theorem. We thus see that, for the proof of claims (a) and (b)–(d)(i), the appropriate sets B and C are of the form $B = (x+1)..N$ and $C = \{x\}$ for $x \in b_*..(N-1)$, whereas for the proof of (b)–(e)(iv) the appropriate sets B and C are of the form $B = 0..(x-1) \cup B_2$ and $C = \{x\}$ for $x \in B_1$. As in 5.2 and 5.4 the most surprising feature of the above result is that the speeds of convergence in the long run do not depend on the exact details of the function b. They only depend on the (relative) position of b_*.

A corollary of the theorem above is the following

6.3 Corollary

Let $b : E \to E$ as in 4.2.

(a) Let $b_* = 0$ then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = 0$.

(b) Let $0 < b_* < \frac{N+1}{3}$. Then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = 0$ and $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = \infty$.

(c) Let $b_* = \frac{N+1}{3}$. Then there exist constants $c, C, k, K > 0$ such that, for sufficiently small ϵ, we have $c \le E^{\gamma^\epsilon}[T_N] \le C$ and $k \le E^{\gamma^\epsilon}[T_{B_2}] \le K$. Moreover, $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = \infty$.

(d) If $\frac{N+1}{3} < b_* < \frac{N+1}{2}$ then

  (i) $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$;

  (ii) $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = \infty$;

  (iii) There exist constants $c, C > 0$ such that $c \le E^{\gamma^\epsilon}[T_{(N-2b_*+1)..N}] \le C$. Moreover, for all $1 \le x \le N - 2b_*$, we have $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{x..N}] = 0$.

(e) If $b_* = \frac{N+1}{2}$ then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$ and $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \infty$. Moreover, there exists a constant $C > 0$ such that, for sufficiently small ϵ, we have $0 \le E^{\gamma^\epsilon}[T_{\{0,N\}}] \le C$.

(f) Let $\frac{N+1}{2} < b_* < \frac{2N+2}{3}$. Then

  (i) $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$.

  (ii) Moreover, $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \infty$.

  (iii) Finally, there exist constants $c, C > 0$ such that, for sufficiently small ϵ, we have $c < E^{\gamma^\epsilon}[T_{0..(2N+1-2b_*)}] < C$. Moreover, for all $2N + 2 - 2b_* \le x \le N - 1$, we have $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{0..x}] = 0$.

(g) Let $b_* = \frac{2N+2}{3}$. Then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$. Moreover, there exist constants $c, C, k, K > 0$ such that

$$c < E^{\gamma^\epsilon}[T_{B_1}] < C \quad \text{and} \quad k < E^{\gamma^\epsilon}[T_0] < K.$$

(h) Let $b_* > \frac{2N+2}{3}$ then $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_N] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$ and $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_0] = \lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_1}] = 0$.

Proof. Claim (a) follows immediately from 6.2(a). The first claim of Part (b) follows immediately from 6.2(b)(i)–(ii) together with the equivalence between the assertions $b_* < \frac{N+1}{3}$ and $N - 3b_* + 1 > 0$ and a simple limit operation. Moreover, since $B_2 \neq \emptyset$, the second claim of this part follows from 6.2(b)(iii)–(iv) together with the inequality $b^* \leq N-1$. The proof of Part (c) follows from similar argumentations as the one used above together with 6.2(c). Claims (d)(i) and (ii) follow immediately from 6.2(c)(i) and (ii) together with the equivalence between the assertions $\frac{N+1}{3} < b_*$ and $3b_* - N - 1 > 0$ and a simple limit operation. In order to prove (d)(iii), we apply result 6.2 (d)(v): Let first $x = N+1-2b_*$. Simple algebraic operations show that $1 \leq N+1-2b_* \leq b^*$ if and only if $\frac{N+2}{3} \leq b_* \leq \frac{N}{2}$. But this last assertion is implied by $\frac{N+1}{3} < b_* < \frac{N+1}{2}$ (In order to see this, discern between even and odd N). Hence, from (d)(v) follows $E^{\gamma^\epsilon}[T_{(N+1-2b_*)..N}] = \mathcal{O}(1)$. This last assertion implies plainly the first claim of (d)(iii). Moreover, simple algebraic manipulations show that

$$b_* > \frac{N+1}{3} \Rightarrow b_* \geq \frac{N+1}{3} \Rightarrow N - 2b_* \leq b^* \Rightarrow 1..(N-2b_*) \subseteq 2..b^*.$$

Therefore, from (d)(v) it follows $E^{\gamma^\epsilon}[T_{x..N}] = \mathcal{O}(\epsilon^{N+1-2b_*-x})$ for all $x \in 1..(N-2b_*)$ and the second claim of (d)(iii) follows from a limit operation. Part (e) follows immediately from 6.2(e) together with the fact $N \geq 2$. The claims (f)–(h) follow from (b)–(d) together with the remark at the end of Subsection 4.1. □

6.4 Remarks

(a) Consider 6.3(d)(iii). We can find a number $x \in 1..(N-2b_*)$ if and only if $\frac{N+1}{3} < b_* \leq \frac{N-1}{2}$. Similarly, in 6.3(f)(iii) the interval $(2N+2-2b_*)..(N-1)$ is nonempty if and only if $\frac{N+1}{2} \leq b_* < \frac{2N+1}{3}$. A necessary condition for both inequalities is $N \geq 6$.

(b) Note that, for all possible values of b_*, there is (at least) one set $\emptyset \neq C \subseteq E$ such that $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_C] \in \{\ 0,\ \text{constant}\ \}$. If $b_* \leq \frac{N+1}{3}$ then $C = \{N\}$. If $\frac{N+1}{3} < b_* < \frac{N+1}{2}$ then $C = (N-2b_*+1)..N$. Finally, if $b_* = \frac{N+1}{2}$ then $C = \{0, N\}$.

(c) Consider the case $\frac{N+1}{3} < b_* < \frac{N+1}{2}$. By the last paragraph in 4.2, we have $\lim_{\epsilon\to 0} \gamma^\epsilon(B_2) = 1$, i.e., for sufficiently small ϵ, the stationary distribution approaches the Dirac measure on N. Therefore, one may expect that $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = 0$, an assertion which is not true, cf. 6.3(d)(i). This fact can be easily explained taking into account (i) that transitions from an element in B_1 to the set B_2 are expensive:

$$Q^\epsilon(x, B_2) = P[S^\epsilon(b(x)) \geq b_*] = \mathcal{O}(\epsilon^{b_* - b(x)}), \qquad x \in B_1,$$

and (ii) the equality

$$E^{\gamma^\epsilon}[T_{B_2}] = \sum_{x\in B_1} \gamma^\epsilon(x) E^x[T_{B_2}].$$

Therefore, since the mean first passage time to the set B_2 involves only transitions of this kind and these transitions rarely occur, the result $\lim_{\epsilon\to 0} E^{\gamma^\epsilon}[T_{B_2}] = \infty$ is not surprising at all. The situation becomes different, if we consider the mean first passage time to the set

$(N-2b_*+1)..N$. Since the value $N-2b_*+1$ is in the domain of attraction of the set B_1, we may write $E^{\gamma^\epsilon}[T_{(N-2b_*+1)..N}]$ as

$$\sum_{x\in 0..(N-2b_*)} \gamma^\epsilon(x) E^x[T_{(N-2\ b_*+1)..N}].$$

This expression involves transitions from an element in the set B_1 to a set including elements of this first one. These transitions are less expensive than those mentioned above:

$$Q^\epsilon(x,(N-2b_*+1)..N) = P[S^\epsilon(b(x)) \geq N-2b_*+1] = \mathcal{O}(\epsilon^{N-2b_*+1-b(x)}),$$

for all $x \in 0..(N-2b_*)$. These two facts explain why $E^{\gamma^\epsilon}[T_{(N-2b_*+1)..N}]$ is bounded.

As in Corollary 5.6, if we take into account Theorem 6.2 and Corollary 6.3 above together with the Remark 4.3(a) we can compute the speed of convergence and the limit of $E^{\gamma^\epsilon}[T_A]$ as ϵ approaches 0. These parameters depend on the entries of the payoff matrix (14).

Acknowledgment. I thank Professor Gunter Ritter for pointing out the topic to me and for reading the manuscript. I would also like to thank the two anonymous referees of an earlier version of this paper for their valuable comments and for pointing reference [7] to me.

A Appendix

Consider the following complement of the Model 4.2. First, there exists an element $\tilde{x} \in 1..(N-1) \neq \frac{N}{2}$ such that $b(\tilde{x}) = \tilde{x}$ and second, instead of (i) and (ii) we have

(i') B_1, $\{\tilde{x}\}$ and B_2 are consecutive intervals, and

(ii') $B_1 \cup \{\tilde{x}\} \cup B_2 = 1..(N-1)$.

Let us denote by $K^\epsilon \in \mathbb{R}^{E\times E}$ the transition matrix generated by b, cf. (18). Similar arguments as the one used in 4.2 show that K^ϵ is irreducible. The aim of this appendix is to derive some convergence statements for the stationary distribution γ^ϵ of the Markov kernel K^ϵ. For this sake, consider the following

A.1 Definitions and Remarks

(a) Let $b^L : E \to E$ be defined by

$$b^L(x) = \begin{cases} 0, & x \in 0..\tilde{b} =: B_1^L; \\ (x+1)\wedge N, & x \in (\tilde{b}+1)..N =: B_2^L. \end{cases}$$

Clearly, the triple (b^L, B_1^L, B_2^L) satisfy the requirements of Model 4.2. Let us denote by $K^{\epsilon,L} \in \mathbb{R}^{E\times E}$ the transition matrix generated by b^L, cf. (18), and by $\gamma^{\epsilon,L}$ its stationary distribution.

(b) Let $b^U : E \to E$ be defined by

$$b^U(x) = \begin{cases} (x-1)\vee 0, & 0..(\tilde{b}-1) =: B_1^U; \\ N, & x \in \tilde{b}..N =: B_2^U. \end{cases}$$

Clearly, the triple (b^U, B_1^U, B_2^U) satisfies the requirements of Model 4.2. Let us denote by $K^{\epsilon,U} \in \mathbb{R}^{E\times E}$ the transition matrix generated by b^U, cf. (18), and by $\gamma^{\epsilon,U}$ its stationary distribution.

(c) Clearly, for the random variables $\mathrm{Bin}(k,\epsilon)$, $k \in 0..N$, we have

$$\mathrm{Bin}(0,\epsilon) \preceq \mathrm{Bin}(1,\epsilon) \preceq \ldots \preceq \mathrm{Bin}(N,\epsilon). \tag{75}$$

(d) Since the function b^L is increasing (75), together with standard arguments on stochastic ordering, implies that the transition matrix $K^{\epsilon,L}$ is stochastically monotone. An analogous argument holds for $K^{\epsilon,U}$. Again, using (75) it is not difficult to see that $K^{\epsilon,L} \preceq K^{\epsilon} \preceq K^{\epsilon,U}$. Therefore,

   (i) $\gamma^{\epsilon,L} \preceq \gamma^{\epsilon} \preceq \gamma^{\epsilon,U}$; and

   (ii) $E^{\gamma^{\epsilon,L}}[T_{x..N}] \leq E^{\gamma^{\epsilon}}[T_{x..N}] \leq E^{\gamma^{\epsilon,U}}[T_{x..N}]$ for all $x \in E$. Furthermore, $E^{\gamma^{\epsilon,L}}[T_{0..y}] \geq E^{\gamma^{\epsilon}}[T_{0..y}] \geq E^{\gamma^{\epsilon,U}}[T_{0..y}]$ for all $y \in E$.

The Item (i) above implies that

$$\text{if } \lim_{\epsilon\to 0} \gamma^{\epsilon,L}(N) = 1, \lim_{\epsilon\to 0} \gamma^{\epsilon}(N) = 1;$$

and

$$\text{if } \lim_{\epsilon\to 0} \gamma^{\epsilon,U}(0) = 1, \lim_{\epsilon\to 0} \gamma^{\epsilon}(0) = 1.$$

With these tools at hand we are now able to state the following complement of Corollary 5.6. Since its proof follows immediately from Theorem 5.4 together with the previous observations, we omit it.

A.2 Corollary (Supplement of Corollary 5.6)

Consider the complement of Model 4.2 given above. If $\tilde{x} < \frac{N-1}{2}$ then $\lim_{\epsilon\to 0} \gamma^{\epsilon}(N) = 1$, whereas if $\tilde{x} > \frac{N+1}{2}$ then $\lim_{\epsilon\to 0} \gamma^{\epsilon}(0) = 1$.

Similarly, this time using A.1(d)(ii) and Corollary 6.3, we can also compute the limits $\lim_{\epsilon\to 0} E^{\gamma^{\epsilon}}[T_A]$, where A is of type $x..N$ or $0..y$, $x \in 1..N$, $y \in 0..(N-1)$.

References

[1] Canals J. and F. Vega–Redondo (1998) Multi–level evolution in population games. International Journal of Game Theory 27:21–35.

[2] Cronshaw M. and L. Kazarian (1997) Some Bounds for Markov Chains (1997). International Journal of Game Theory 26:579–582.

[3] Harsanyi, J., and R. Selten (1988) A General Theory of Equilibrium in Games. Cambridge: MIT Press

[4] Ellison G. (1993) Learning, local interaction, and coordination. Econometrica 61: 1047–1072.

[5] Kandori M., G.J. Mailath, and R. Rob (1993) Learning, Mutations and Long Run Equilibria in Games. Econometrica 61: 29–56.

[6] Kemeny, J. and L. Snell (1976) Finite Markov Chains. Springer–Verlag, New York.

[7] Rhode, P., and M. Stegeman (1996) A Comment on Learning, Mutation, and Long Run Equilibria in Games. Econometrica 64: 443-449

[8] Stoyan, D. (1993), Comparison Methods for queues and other stochastic methods. John Wiley.

[9] Young, P. (1993), The Evolution of Conventions. Econometrica 61: 57–84.