OPTIMAL STOPPING AND STRONG APPROXIMATION THEOREMS.

YURI KIFER

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL

Abstract. Strong approximation theorems known also as (strong) invariance principles provide uniform (in time) almost sure or in average approximations (as opposed to the convergence in distribution) in the central limit theorem type results which is done by redefining in certain ways corresponding random variables or vectors on one probability space without changing their distributions. Three methods are known at present to provide appropriate constructions and they yield also estimates of approximation errors. In this paper we are interested in error estimates for approximations of values of Dynkin's optimal stopping games with payoff processes which are Lipschitz functionals of the Brownian motion by values of sequences of Dynkin's games with payoff processes converging to payoffs of original games. The strong approximations described above play a crucial role here but we have to face an additional substantial complication to have all stopping times defined with respect to the same filtration. The Skorokhod embedding method employed in [10] is well suited for this purpose but it does not work in the multidimensional case where another method from [1] should be applied. The most precise quantile method of strong approximations does not seem to work for this type of problems and the question about optimality of error estimates is not clear yet. The results are new for convergence of corresponding Snell's envelopes, as well, and they are motivated, in particular, by financial mathematics applications of approximations of fair prices of American and game (Israeli) options introduced in [8].

1. Introduction

In the last 40 years starting with Strassen's work (see [22]) strong approximation theorems attracted substantial attention in probability (see [3] and references there). The approach was based at the beginning on the Skorokhod embedding of martingales into the Brownian motion. Later it was understood that this method does not provide optimal error terms of approximations and a more efficient quantile transform method was created (see [11] and [24]). Still, both methods have their limitations: Skorokhod embeddings work only in the one dimensional case and quantile transforms are useful only for sums of independent random variables and vectors. Another method developed in [1] and based on approximations of

Date: August 28, 2006.
2000 *Mathematics Subject Classification.* Primary 91B28: Secondary: 60F15, 91A05.
Key words and phrases. Dynkin games, game options, strong invariance principles.
Partially supported by an ISF grant.

random variables with close characteristic functions works both in the multidimensional case and for sums of weakly dependent random variables and vectors though it does not lead to approximations with optimal error estimates.

Discrete time approximations in optimal stopping problems have their few decades history too but most of the work was done in the case of payoff functions dependent only on the last state of an underlying Markov process which led usually to difference schemes approximations of second order elliptic or parabolic differential equations (see [13]). When payoffs depend on the whole path of an underlying process such methods do not work and other approaches should be considered. Some general results on weak convergence of Snell's envelopes assuming weak convergence of underlying processes were obtained in [14] and [16] but the conditions required for their approach are difficult to check except for Markov payoffs and it does not work in a more or less general path dependent case. Moreover, their machinery cannot provide, in principle, speed of convergence estimates. Thus methods based on strong approximation theorems come into the picture.

Among motivations for the study of the above convergence are computational problems in financial mathematics where the fair price of American or, more generally, game (Israeli) options (the latter introduced in [8]) cannot be represented by an explicit formula and for path dependent payoffs in the Black-Scholes markets we also cannot employ partial differential equations methods. The only available strategy here is to approximate (in a strong sense) Black-Scholes markets by a sequence of Cox-Ross-Rubinstein binomial models and to show that prices of the corresponding American (Snell's envelopes) and game options converge, as well, with error estimates coming along (see [10]).

In this paper we start with a strong approximation of the Brownian motion by means of properly normalized sums of independent or weakly dependent random vectors, then consider optimal stopping (Dynkin) games with payoffs depending on these sums and, finally, prove that under appropriate conditions values of the games with payoffs based on these sums converge to the value of the game whose payoff is based on the Brownian motion and the error estimates will be given, as well. In order to compare the corresponding prices we have not only to consider all relevant processes on one probability space, which is standard in strong approximation theorems, but also to ensure that all stopping times are defined with respect to the same filtration. We note that the quantile transform method led recently to essentially optimal error estimates in strong approximation theorems for sums of independent random vectors (see [24]) but this approach does not seem to be applicable to our problems since constructions there do not respect filtrations in a very substantial way as they depend on the far away future. We have not made special efforts to obtain sharpest possible estimates even within our method since it itself does not seem to be optimal and the question about optimality of error estimates of such approximations in the optimal stopping framework is not clear yet.

2. Preliminaries and main results

Our setup includes ψ-mixing sequences of random vectors so we start with the corresponding definitions. Given a probability space $(\Omega, \mathcal{F}, P)$ the ψ-mixing coefficient of two σ-subalgebras $\mathcal{G}, \mathcal{H} \subset \mathcal{F}$ is defined by (see [4]),

$$\psi(\mathcal{G}, \mathcal{H}) = \sup\big\{\big|1 - \frac{P(\Gamma \cap H)}{P(\Gamma)P(H)}\big| : \Gamma \in \mathcal{G},\ P(\Gamma) > 0,\ H \in \mathcal{H},\ P(H) > 0\big\}. \tag{2.1}$$

Let $X_1, X_2, ..., EX_j = 0$ be a stationary centered sequence of random vectors having values in the d-dimensional Euclidean space $\mathbb{R}^d$. Denote by $\mathcal{F}_{n,m}$, $m \geq n$ the σ-subalgebra of $\mathcal{F}$ generated by $\{X_j,\ n \leq j < m\}$. The sequence $X_1, X_2, ...$ is called ψ-mixing if

$$\psi(m) = \sup_{n \geq 1} \psi(\mathcal{F}_{1,n}, \mathcal{F}_{n+m,\infty}) \downarrow 0 \text{ as } m \uparrow \infty. \tag{2.2}$$

Among prominent ψ-mixing (but not independent) stationary sequences we can mention the sequence of digits of continued fractions of points of the unit interval considered with the Gauss measure where the mixing coefficient $\psi(m)$ decays exponentially fast in $m \to \infty$ (see [6] and references there).

Consider normalized sums defined for all $t \in [0, T]$, $T > 0$ by

$$S_n^X(t) = (T/n)^{1/2} \sum_{i=1}^{[tn/T]} X_i \text{ if } t \geq T/n \text{ and } S_n^X(t) = 0, \text{ otherwise.} \tag{2.3}$$

Often it is more convenient to consider continuous in time processes interpolating linearly between values $S_n^X(kT/n)$ and $S_n^X((k+1)T/n)$ but in our setup the above piecewise constant process is more natural to start with. Let $X_k = (x_1^{(k)}, ..., x_d^{(k)})$, $k = 1, 2, ...$ Under even weaker than ψ-mixing conditions but assuming that the mixing coefficient decays fast enough for any $i, j = 1, ..., d$ the two series in

$$a_{ij} = Ex_i^{(1)} x_j^{(1)} + \sum_{k \geq 2} Ex_i^{(1)} x_j^{(k)} + \sum_{k \geq 2} Ex_i^{(k)} x_j^{(1)} \tag{2.4}$$

converge absolutely and the invariance principle type limit theorems say that the process $S_n^X(t)$, $t \in [0, T]$ converges weakly to the d-dimensional Brownian motion $B_A(t)$, $t \in [0, T]$ with the covariance matrix $A = (a_{ij})$, i.e. $B_A(t)$ is a d-dimensional (Gaussian) process with independent increments starting at zero and such that $B_A(t) - B_A(s)$ has normal distribution with mean 0 and the covariance matrix $(t - s)A$. Strong approximation theorems (see [1] and [3]) assert that, moreover, we can consider both S_n^X and B_A on one probability space and obtain almost sure or in average estimates for $\sup_{0 \leq t \leq T} |S_n^X(t) - B_A(t)|$. Observe that if $X_1, X_2, ...$ are independent then the matrix A above coincides with the covariance matrix of X_1.

Now let $F_t \leq G_t$ be continuous functionals on the space $\mathbb{M}_d[0, t]$ of bounded Borel measurable maps from $[0, t]$ to $\mathbb{R}^d$ (actually, we will deal only with piecewise constant and continuous maps) considered with the uniform metric $d_{0t}(\upsilon, \tilde{\upsilon}) = \sup_{0 \leq s \leq t} |\upsilon_s - \tilde{\upsilon}_s|$. Consider an optimal stopping (Dynkin) game of two players **I** and **II** where **I** pays to **II** the amount $F_t(S_n^X)$ or $G_t(S_n^X)$ if **I** or **II** stops the game at time t, respectively. If nobody stops the game before the (horizon) time T then it is stopped automatically at T in which case **I** pays to **II** the amount $F_T(S_n^X)$.

Set

$$R_n^X(s,t)=G_s(S_n^X)\mathbb{I}_{s<t}+F_t(S_n^X)\mathbb{I}_{t\le s}, \tag{2.5}$$

where $\mathbb{I}_\Gamma=1$ if Γ occurs and 0, otherwise. Then the value of the above game can be obtained by the formula (see [17] and [15]),

$$V_n^X=\inf_{\zeta\in\mathcal{T}_n^X}\sup_{\eta\in\mathcal{T}_n^X}ER_n^x(\zeta,\eta) \tag{2.6}$$

where $\mathcal{T}_n^X$ is the space of stopping times with respect to the filtration $\mathcal{F}_0^X=\{\emptyset,\Omega\}$, $\mathcal{F}_k^X=\mathcal{F}_{1,k+1}$, $k=1,2,...,n$ taking on values in the set $\{0,T/n,2T/n,...,(n-1)T/n,T\}$. A corresponding game based on the Brownian motion $B_A(t)$, $t\in[0,T]$ in place of sums S_n^X with the payoffs

$$R^B(s,t)=G_t(B_A)\mathbb{I}_{s<t}+F_t(B_A)\mathbb{I}_{t\le s} \tag{2.7}$$

has the value

$$V^B=\inf_{\sigma\in\mathcal{T}_{0T}^B}\sup_{\tau\in\mathcal{T}_{0T}^B}ER^B(\sigma,\tau) \tag{2.8}$$

where $\mathcal{T}_{0T}^B$ is the space of stopping times with respect to the Brownian filtration taking on values in $[0,T]$. In this paper we will be interested to establish convergence $V_n^X\to V^B$ as $n\to\infty$ and to estimate its speed.

We will work mostly with functionals F and G for which there exists a constant $L\ge 1$ such that for any $t\ge s\ge 0$, $s,t\in[0,T]$ and $\upsilon,\tilde\upsilon\in\mathbb{M}_d[0,t]$,

$$|F_t(\upsilon)-F_t(\tilde\upsilon)|+|G_t(\upsilon)-G_t(\tilde\upsilon)|\le Ld_{0t}(\upsilon,\tilde\upsilon) \tag{2.9}$$

and

$$|F_t(\upsilon)-F_s(\upsilon)|+|G_t(\upsilon)-G_s(\upsilon)|\le L\big(|t-s|(1+\sup_{u\in[s,t]}|\upsilon_u|)+\sup_{u\in[s,t]}|\upsilon_u-\upsilon_s|\big). \tag{2.10}$$

Exponential functionals which emerge usually in financial mathematics formulas for prices of game options (see [8] and [10]) do not satisfy the Lipshitz condition (2.9) and the assumption (2.10). In order to include some exponential functionals we will consider also less restrictive than (2.9) and (2.10) conditions where for some $L\ge 1$ and all $T\ge t\ge s\ge 0$, and $\upsilon,\tilde\upsilon\in\mathbb{M}_d[0,t]$,

$$\begin{aligned}|F_t(\upsilon)-F_t(\tilde\upsilon)|+|G_t(\upsilon)-G_t(\tilde\upsilon)|\le L\big(d_{0t}(\upsilon,\tilde\upsilon)+\mathbb{I}_{\sup_{0\le u\le t}|\upsilon_u-\tilde\upsilon_u|>1}\big)\\ \times\exp\big(L\sup_{0\le u\le t}|\upsilon_u|+L\sup_{0\le u\le t}|\tilde\upsilon_u|\big)\end{aligned} \tag{2.11}$$

and

$$|F_t(\upsilon)-F_s(\upsilon)|+|G_t(\upsilon)-G_s(\upsilon)|\le L\big(|t-s|+\sup_{u\in[s,t]}|\upsilon_u-\upsilon_s|\big)\exp\big(Lsup_{0\le u\le t}|\upsilon_u|\big). \tag{2.12}$$

2.1. **Theorem.** *Let $X_1,X_2,...,$ $EX_1=0$ be a ψ-mixing stationary centered sequence of $\mathbb{R}^d$-valued random vectors such that*

$$E|X_j|^4<\infty\quad\textit{and}\quad\sup_{k\ge1}k^4\psi(k)<\infty. \tag{2.13}$$

Then the two series in (2.4) converge absolutely and if the functionals F_t and G_t, $t\in[0,T]$ satisfy conditions (2.9) and (2.10) then there exists a constant $C=C_T>0$ such that for all $n\ge1$,

$$|V_n^X-V^B|\le Cn^{-1/(26556+13440d)}\sqrt{\log n} \tag{2.14}$$

where V_n^X *and* V^B *are defined by (2.5)–(2.8). Assume that with probability one*

$$|X_1| \le \tilde{C} \tag{2.15}$$

for some constant $\tilde{C} > 0$ *then (2.14) still holds true (with may be an extra factor* $\sqrt{\log n}$*) under the conditions (2.11) and (2.12) in place of (2.9) and (2.10). If* $X_1, X_2, ...$ *are independent with a nondegenerate covariance matrix then the right hand side of the estimate (2.14) can be improved to* $n^{-1/96(1+d)} \log n$*.*

We observe that the conditions of Theorem 2.1 are more than enough in order to use the results of [17] and [15] concerning existence of values of the above Dynkin games given by (2.6) and (2.8). Note also that the condition (2.13) can be relaxed somewhat but the estimate (2.14) depends on the form of this assumption and we prefer to choose a specific form of it though we cannot provide a sharp estimate in (2.14). The main bottleneck for a substantial improvement of the estimate (2.14) is the use of Theorem 3.23 from [3] (see Lemma 4.1 below) which estimates characteristic functions of normalized sums of mixing sequences of random vectors. In the case of independent random vectors much better estimates of characteristic functions of their sums can be given which leads to a better estimate in (2.14), as well. In principle, the constant C in (2.14) can be estimated from the proof and it depends on the bounds in (2.13). Among functionals satisfying the conditions (2.9) and (2.10) we can mention $F_t(\upsilon) = \int_0^t f(\upsilon_s)ds$ and $F_t(\upsilon) = \sup_{0\le s\le t} f(\upsilon_s)$ for some Lipschitz continuous function f on $\mathbb{R}^d$ and, on the other hand, the functionals $F_t(\upsilon) = \int_0^t e^{f(\upsilon_s)}ds$ and $F_t(\upsilon) = \sup_{0\le s\le t} e^{f(\upsilon_s)}$ satisfy the conditions (2.11) and (2.12). Observe that estimates of Theorem 2.1 are new not only for optimal stopping games but also for Snell's envelopes, i.e. for optimal stopping by one player.

The strategy of the proof of Theorem 2.1 consists of the following. First, we coarsen the time allowing only stopping times which take on values $0, t_1(n), t_2(n), ..., t_{\ell(n)}(n)$ where $t_k(n) = kmT/n$ for some $m = M + N$ with $M, N \in \mathbb{N}$ and $\ell(n) = [n/m]$. Here M is large but still much smaller than N which will be used on the second step. This changes game values by certain amount which will be estimated in Section 5. On the first step we do not change the filtrations yet. On the second step we split the whole normalized sum S_n^X into the sum of blocks $U_k^{(n)} = \sum_{mk\ge j\ge 1+m(k-1)} X_j$ which are weakly dependent if we disregard small initial subblocks of length M whose contribution can be estimated separately. Now we pass to the filtration $\{\mathcal{F}_k^U,\, k = 1, ..., \ell(n)\}$ generated by $U_j^{(n)}$, $j = 1, ..., k$ and allow only stopping times with respect to this new filtration of smaller σ-algebras. Relying on ψ-mixing we are able to estimate again (in Section 3) the corresponding change in the game value. We do a similar modification with the game whose payoff is based on the Brownian motion B_A with the covariance matrix A by considering a subfiltration $\mathcal{G}_k^B$, $k = 1, ..., \ell(n)$ generated by increments $B_A(t_j(n)) - B_A(t_{j-1}(n))$, $j = 1, ..., k$. This does not lead, in fact, to a change in the game value (see Lemma 3.2). In Section 4 we employ the strong approximation theorem from [3] in order to compare game values based on sums and on B_A obtained after two steps above. Now we are able not only to consider both processes on the same probability space but also to deal with stopping times defined with respect to the same filtration. Finally, in Section 5 we derive remaining estimates mentioned at the beginning of this description and complete the proof of Theorem 2.1 by collecting all estimates and choosing suitable parameters of our constructions.

3. Key lemmas

We start with the following important result.

3.1. Lemma. *Let $h=h(x,y)$, $x\in\mathbb{R}^l, y\in\mathbb{R}^m$ be a Borel measurable function on $\mathbb{R}^l\times\mathbb{R}^m$, $(\Omega,\mathcal{F},P)$ be a probability space, $\mathcal{G},\mathcal{H}\subset\mathcal{F}$ be σ-subalgebras, $X:\Omega\to\mathbb{R}^l$ and $Y:\Omega\to\mathbb{R}^m$ be random vectors such that X is $\mathcal{G}$-measurable and Y is $\mathcal{H}$-measurable. Suppose that*

$$r(x)=E|h(x,Y)|<\infty,\ \forall x\in\mathbb{R}^l \quad \textit{and} \quad E|h(X,Y)|<\infty, \tag{3.1}$$

and set $g(x)=Eh(x,Y)$. Then

$$\big|E\big(h(X,Y)\big|\mathcal{G}\big)-g(X)\big|\le\psi(\mathcal{G},\mathcal{H})r(X), \tag{3.2}$$

where $\psi(\mathcal{G},\mathcal{H})$ is the same as in (2.1), and (3.2) does not hold true, in general, with a smaller than $\psi(\mathcal{G},\mathcal{H})$ coefficient.

Proof. The assertion is well known when $\psi(\mathcal{G},\mathcal{H})=0$, i.e. when Y is independent of $\mathcal{G}$ (see, for instance, [2], Example 34.3) and a particular case of (3.2) for $h(x,y)=y\in\mathbb{R}^1$ appears as Lemma 1.2 in [18]. Still, since we could not find our more general statement in the literature, the detailed argument follows.

Let μ_X,μ_Y, and $\mu_{X,Y}$ be the probability measures which are the images of P under the maps X, Y, and $(X,Y):\Omega\to\mathbb{R}^l\times\mathbb{R}^m$, respectively, and set $\mu=\frac{1}{2}(\mu_X\times\mu_Y+\mu_{X,Y})$. Now observe that it suffices to prove (3.2) only for functions h which are continuous and have compact supports since this set of functions is dense in the space $\mathbb{L}^1(\mathbb{R}^l\times\mathbb{R}^m,\mu)$ (see, for instance, [20], Theorem 3.14). Next, let h be a continuous function on $\mathbb{R}^l\times\mathbb{R}^m$ whose support is contained in $K=K^{(l)}\times K^{(m)}$ where, say, $K^{(l)}=\{x\in\mathbb{R}^l:|x|\le M\}$ and $K^{(m)}=\{x\in\mathbb{R}^m:|x|\le M\}$ for some $M>0$. Then h is uniformly continuous in K, and so given $\varepsilon>0$ there exists $\delta_\varepsilon>0$ such that $|h(x,y)-h(\tilde x,\tilde y)|\le\varepsilon$ whenever $(x,y),(\tilde x,\tilde y)\in K$ and $|x-\tilde x|\le\delta_\varepsilon$ and $|y-\tilde y|\le\delta_\varepsilon$. Choose $\delta_\varepsilon-$nets $\{x_1,...,x_{n_\varepsilon}\}$ and $\{y_1,...,y_{n_\varepsilon}\}$ in $K^{(l)}$ and $K^{(m)}$, respectively, and Borel measurable partitions $\{U_1,...,U_{n_\varepsilon}\}$ and $\{V_1,...,V_{n_\varepsilon}\}$ of $K^{(l)}$ and $K^{(m)}$, respectively, such that $x_i\in U_i, y_j\in V_j$ and $|x-x_i|\le\delta_\varepsilon$, $|y-y_j|\le\delta_\varepsilon$ for any $x\in U_i$ and $y\in V_j$, $i,j=1,...,n_\varepsilon$. Set $A_i=\{\omega\in\Omega:\ X(\omega)\in U_i\}$ and $B_i=\{\omega\in\Omega:\ Y(\omega)\in V_i\}$, $i=1,...,n_\varepsilon$, Then

$$|h(X,Y)-\sum_{i,j=1}^{n_\varepsilon}h(x_i,y_j)\mathbb{I}_{A_i}\mathbb{I}_{B_j}|\le\varepsilon, \tag{3.3}$$

$$|g(X)-\sum_{i,j=1}^{n_\varepsilon}h(x_i,y_j)\mathbb{I}_{A_i}P(B_j)|\le 2\varepsilon, \tag{3.4}$$

and

$$|r(X)-\sum_{i,j=1}^{n_\varepsilon}|h(x_i,y_j)|\mathbb{I}_{A_i}P(B_j)|\le 2\varepsilon. \tag{3.5}$$

Observe, that for any $D\in\mathcal{G}$ and $B\in\mathcal{H}$ with $P(D)>0$ and $P(B)>0$,

$$\frac{1}{P(D)}\int_D\big(\frac{P(B|\mathcal{G})}{P(B)}-1\big)dP=\frac{P(B\cap D)}{P(B)P(D)}-1.$$

Taking the absolute value of the integrand and then the essential supremum over all positive probability events $B \in \mathcal{H}$ we obtain that with probability one

$$esssup_{B\in\mathcal{H},P(B)>0}\Big|\frac{P(B|\mathcal{G})}{P(B)}-1\Big| \geq \psi(\mathcal{G},\mathcal{H}).$$

This inequality shows that already in the particular case $h(x,y)=y$ and $Y=\mathbb{I}_B$, $B\in\mathcal{H}$ we cannot obtain, in general, (3.2) with smaller than $\psi(\mathcal{G},\mathcal{H})$ coefficient (which, essentially, was indicated to me by M.Peligrad). In fact, we will need below the inequality in the other direction which will say that, actually, we have an equality here. Namely, since the integrand above is $\mathcal{G}$-measurable we can write the above equality for $D^{(+)}$ where the integrand is positive and for $D^{(-)}$ where it is negative concluding that for any positive probability events $D\in\mathcal{G}$ and $B\in\mathcal{H}$,

$$\frac{1}{P(D)}\int_D\Big|\frac{P(B|\mathcal{G})}{P(B)}-1\Big|dP \leq \psi(\mathcal{G},\mathcal{H}).$$

Again, by $\mathcal{G}$-measurability of the integrand it follows that with probability one

$$esssup_{B\in\mathcal{H},P(B)>0}\Big|\frac{P(B|\mathcal{G})}{P(B)}-1\Big| \leq \psi(\mathcal{G},\mathcal{H}).$$

Now, this together with (3.5) yield that with probability one,

$$\begin{aligned}\big|E\big(\textstyle\sum_{i,j=1}^{n_\varepsilon} h(x_i,y_j)\mathbb{I}_{A_i}\mathbb{I}_{B_j}\big|\mathcal{G}\big)-\textstyle\sum_{i,j=1}^{n_\varepsilon} h(x_i,y_j)\mathbb{I}_{A_i}P(B_j)\big|\\ \leq \textstyle\sum_{i,j=1}^{n_\varepsilon}|h(x_i,y_j)|\mathbb{I}_{A_i}|P(B_j|\mathcal{G})-P(B_j)|\\ \leq \psi(\mathcal{G},\mathcal{H})\textstyle\sum_{i,j=1}^{n_\varepsilon}|h(x_i,y_j)|\mathbb{I}_{A_i}P(B_j)\leq\psi(\mathcal{G},\mathcal{H})(r(X)+2\varepsilon).\end{aligned}$$

This together with (3.3) and (3.4) yield (3.2) by letting $\varepsilon\to 0$. □

Next, our constructions will depend on two large integers N and M such that N is much smaller than n and M is much smaller than N but their precise choice will be specified later on. Set $\ell(n)=[\frac{n}{N+M}]$ and split the sum of random vectors $X_1,...,X_n$ into blocks $U_k^{(n)}=Y_k^{(n)}+Z_k^{(n)}$ with the remainder $Q^{(n)}=\sum_{i=\ell(n)(M+N)+1}^{n}X_i$, where for $k=1,2,...,\ell(n)$,

$$Y_k^{(n)}=\sum_{i=1}^{N}X_{i+kM+(k-1)N} \text{ and } Z_k^{(n)}=\sum_{i=1}^{M}X_{i+(k-1)(M+N)}.$$

Our choice of M and N will make $Y_k^{(n)}$ "almost independent" of $\mathcal{F}^X_{(k-1)(M+N)}$ while keeping the total contribution of $Z_k^{(n)}$, $k=1,...,\ell(n)$ sufficiently small. Set $t_k(n)=kT(N+M)/n$, $k(t)=[\frac{tn}{T(N+M)}]$ and

$$S_n^U(t)=(T/n)^{1/2}\sum_{j=1}^{k(t)}U_j^{(n)} \text{ if } t\geq t_1(n) \text{ and } S_n^U(t)=0, \text{ otherwise.} \tag{3.6}$$

Let $\mathcal{F}_k^U$ be the σ-subalgebra of $\mathcal{F}_{1,k(M+N)+1}$ generated by $U_1^{(n)},...,U_k^{(n)}$ and $\mathcal{F}_0^U=\{\emptyset,\Omega\}$. Denote by $\mathcal{S}_n^X$ and $\mathcal{T}_n^U$ the spaces of stopping times with respect to the filtrations $\mathcal{F}^X_{k(M+N)}$ and $\mathcal{F}^U_k$, $k=0,1,...,\ell(n)$, respectively, taking on values $0,t_1(n),t_2(n),...,t_{\ell(n)}(n)$. Let $\mathcal{F}_t^B$, $t\geq 0$ be the filtration generated by the process B_A and denote by $\mathcal{G}_k^B$ the σ-subalgebra of $\mathcal{F}^B_{t_k(n)}$ generated by $\{B_A(t_1(n)),B_A(t_2(n)),...,B_A(t_k(n))\}$. Let $\mathcal{S}_n^B$ and $\mathcal{T}_n^B$ be the spaces of stopping times with respect to the filtrations $\mathcal{F}^B_{t_k(n)}$ and $\mathcal{G}^B_k$, $k=0,1,...,\ell(n)$, respectively, taking on values $0,t_1(n),...,t_{\ell(n)}(n)$. Set

$$R^U_n(s,t)=G_t(S^U_n)\mathbb{I}_{s<t}+F_t(S^U_n)\mathbb{I}_{t\leq s}, \tag{3.7}$$

$$V^X_{\mathcal{S}}=\inf_{\zeta\in\mathcal{S}^X_n}\sup_{\eta\in\mathcal{S}^X_n}ER^U_n(\zeta,\eta),\quad V^U_n=\inf_{\zeta\in\mathcal{T}^U_n}\sup_{\eta\in\mathcal{T}^U_n}ER^U_n(\zeta,\eta), \tag{3.8}$$

$$V^B_{\mathcal{S}}=\inf_{\sigma\in\mathcal{S}^B_n}\sup_{\tau\in\mathcal{S}^B_n}ER^B(\sigma,\tau)\text{ and }V^B_n=\inf_{\sigma\in\mathcal{T}^B_n}\sup_{\tau\in\mathcal{T}^B_n}ER^B(\sigma,\tau). \tag{3.9}$$

3.2**. Lemma.** *For all N,M and n as above*

$$\begin{aligned}|V^X_{\mathcal{S}}-V^U_n|\leq 2\psi(M)\ell(n)\big(LT+|F_0(0)|+|G_0(0)|\\ +(2+L)(Tn)^{1/2}E|X_1|\big)+2M\ell(n)(T/n)^{1/2}\big(L+\ell(n)E|X_1|\big)\end{aligned} \tag{3.10}$$

and

$$V^B_{\mathcal{S}}=V^B_n. \tag{3.11}$$

Proof. It is well known and easy to see that both $V^X_{\mathcal{S}}$ and V^U_n given by (3.8) can be obtained by the backward induction (dynamical programming algorithm) so that $V^X_{\mathcal{S}}=V^X_{\mathcal{S},0}$ and $V^U_n=V^U_{n,0}$ where

$$V^X_{\mathcal{S},\ell(n)}=V^U_{n,\ell(n)}=F_T(S^U_n) \tag{3.12}$$

and recursively

$$V^X_{\mathcal{S},k}=\min\Big(G_{t_k(n)}(S^U_n),\max\big(F_{t_k(n)}(S^U_n),E\big(V^X_{\mathcal{S},k+1}|\mathcal{F}^X_{k(N+M)}\big)\big)\Big) \tag{3.13}$$

and

$$V^U_{n,k}=\min\Big(G_{t_k(n)}(S^U_n),\max\big(F_{t_k(n)}(S^U_n),E\big(V^U_{n,k+1}|\mathcal{F}^U_k\big)\big)\Big). \tag{3.14}$$

For any vectors $x_1,x_2,...,x_{\ell(n)}\in\mathbb{R}^d$ set $x^U(t)=\sum_{j=1}^{k(t)}x_j$ where $k(t)$ is the same as in (3.6). In view of (2.9) we can write

$$F_t(x^U)=q_{k(t)}(x_1,...,x_{k(t)})\text{ and }G_t(x^U)=r_{k(t)}(x_1,...,x_{k(t)}) \tag{3.15}$$

where q_k and r_k depend only on $x_1,...,x_k$ and satisfy

$$\max\big(|q_k(x_1,...,x_k)-q_k(\tilde x_1,...,\tilde x_k)|,\ |r_k(x_1,...,x_k)-r_k(\tilde x_1,...,\tilde x_k)|\big)\leq L\sum_{j=1}^k|x_j-\tilde x_j|. \tag{3.16}$$

Next, we show by the backward induction that there exist functions $\Phi_k(x_1,...,x_k)$, $k=1,...,\ell(n)$ and a constant Φ_0 such that

$$|\Phi_k(x_1,...,x_k)-\Phi_k(\tilde x_1,...,\tilde x_k)|\leq L\sum_{j=1}^k|x_j-\tilde x_j|, \tag{3.17}$$

$$\begin{aligned}W^X_k\overset{\text{def}}{=}\big|V^X_{\mathcal{S},k}-\Phi_k\big((T/n)^{1/2}U^{(n)}_1,...,(T/n)^{1/2}U^{(n)}_k\big)\big|\\ \leq\psi(M)\ell(n)\big(LT+|F_0(0)|+|G_0(0)|+2(Tn)^{1/2}E|X_1|\big)\\ +2\ell(n)(T/n)^{1/2}\big(E\big(\textstyle\sum_{j=k+1}^{\ell(n)}|Z^{(n)}_j||\mathcal{F}^X_{k(M+N)}\big)+LME|X_1|\big)\\ +\psi(M)L\ell(n)(T/n)^{1/2}\big(\textstyle\sum_{j=1}^k|U^{(n)}_j|+E\big(\sum_{j=k+1}^{\ell(n)}|U^{(n)}_j||\mathcal{F}^X_{k(M+N)}\big)\big)\end{aligned} \tag{3.18}$$

and

$$W_k^U \stackrel{\text{def}}{=} \big|V_{n,k}^U - \Phi_k\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_k^{(n)}\big)\big|$$
$$\le \psi(M)\ell(n)(LT + |F_0(0)| + |G_0(0)| + 2(Tn)^{1/2}E|X_1|\big)$$
$$+2\ell(n)(T/n)^{1/2}\big(E\big(\sum_{j=k+1}^{\ell(n)}|Z_j^{(n)}||\mathcal{F}_k^U\big) + LME|X_1|\big)$$
$$+\psi(M)L\ell(n)(T/n)^{1/2}\big(\sum_{j=1}^{k}|U_j^{(n)}| + E\big(\sum_{j=k+1}^{\ell(n)}|U_j^{(n)}||\mathcal{F}_k^U\big)\big) \tag{3.19}$$

for all $k = \ell(n), \ell(n) - 1, ..., 1, 0$.

Indeed, (3.17)–(3.19) follow for $k = \ell(n)$ from (3.12), (3.15) and (3.16) taking $\Phi_{\ell(n)}(x_1, ..., x_{\ell(n)}) = q_{\ell(n)}(x_1, ..., x_{\ell(n)})$. Suppose that we already constructed functions Φ_k satisfying (3.17)–(3.19) for all $k \ge l + 1$. Now, set

$$\Phi_l(x_1, ..., x_l) = \min\big(r_l(x_1, ..., x_l), \max\big(q_l(x_1, ..., x_l), h_l(x_1, ..., x_l)\big)\big) \tag{3.20}$$

where $h_l(x_1, ..., x_l) = E\Phi_{l+1}\big(x_1, ..., x_l, (T/n)^{1/2}U_{l+1}^{(n)}\big)$. Since for any $a, \tilde a, b, \tilde b, c, \tilde c$,

$$|\min(a, \max(b, c)) - \min(\tilde a, \max(\tilde b, \tilde c))| \le \max(|a - \tilde a|, |b - \tilde b|, |c - \tilde c|) \tag{3.21}$$

then by (3.16), (3.20) and the induction hypotesis that (3.17) is satisfied for $k = l+1$ we obtain that (3.17) holds true also for $k = l$. Next, we obtain from (3.13), (3.14) and (3.21) that

$$W_l^X \le E\big(W_{l+1}^X|\mathcal{F}_{l(M+N)}^X\big) + \big|E\big(\Phi_{l+1}\big((T/n)^{1/2}U_1^{(n)},$$
$$..., (T/n)^{1/2}U_{l+1}^{(n)}\big)\big|\mathcal{F}_{l(M+N)}^X\big) - h_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big)\big| \tag{3.22}$$

and

$$W_l^U \le E\big(W_{l+1}^U|\mathcal{F}_l^U\big) + \big|E\big(\Phi_{l+1}\big((T/n)^{1/2}U_1^{(n)},$$
$$..., (T/n)^{1/2}U_{l+1}^{(n)}\big)\big|\mathcal{F}_l^U\big) - h_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big)\big|. \tag{3.23}$$

We estimate the first term in the right hand side of (3.22) and (3.23) by the induction hypotesis. In order to estimate the second term there we obtain from (3.17) for $k = l + 1$ that for any $x_1, ..., x_l$,

$$\big|\Phi_{l+1}\big(x_1, ..., x_l, (T/n)^{1/2}U_{l+1}^{(n)}\big) - \Phi_{l+1}\big(x_1, ..., x_l, (T/n)^{1/2}Y_{l+1}^{(n)}\big) \le L(T/n)^{1/2}|Z_{l+1}^{(n)}|, \tag{3.24}$$

and so for

$$g_l(x_1, ..., x_l) = E\Phi_{l+1}\big(x_1, ..., x_l, (T/n)^{1/2}Y_{l+1}^{(n)}\big),$$

$$|h_l(x_1, ..., x_l) - g_l(x_1, ..., x_l)| \le L(T/n)^{1/2}E|Z_{l+1}^{(n)}|. \tag{3.25}$$

On the other hand, by Lemma 3.1 with $\mathcal{G} = \mathcal{F}_{l(M+N)}^X$ and $\mathcal{H} = \mathcal{F}_{l(M+N)+M,\infty}$,

$$\big|E\big(\Phi_{l+1}\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}, (T/n)^{1/2}Y_{l+1}^{(n)}\big)\big|\mathcal{F}_{l(M+N)}^X\big)$$
$$-g_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big)\big|$$
$$\le \psi(M)r_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big) \tag{3.26}$$

and

$$\big|E\big(\Phi_{l+1}\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}, (T/n)^{1/2}Y_{l+1}^{(n)}\big)\big|\mathcal{F}_l^U\big)$$
$$-g_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big)\big|$$
$$\le \psi(M)r_l\big((T/n)^{1/2}U_1^{(n)}, ..., (T/n)^{1/2}U_l^{(n)}\big) \tag{3.27}$$

where $r_l(x_1,...,x_l)=E\big|\Phi_{l+1}\big(x_1,...,x_l,(T/n)^{1/2}Y^{(n)}_{l+1}\big)\big|$. Since $X_1,X_2,...$ is a stationary sequence we obtain that for each $l=0,1,...,\ell(n)-1$,

$$E|Z^{(n)}_{l+1}|\leq ME|X_1| \tag{3.28}$$

and since Φ_{l+1} satisfies (3.17) it follows that

$$\begin{aligned} &r_l\big((T/n)^{1/2}U^{(n)}_1,...,(T/n)^{1/2}U^{(n)}_l\big)\\ &\leq|\Phi_{l+1}(0,...,0)|+L(T/n)^{1/2}\big(NE|X_1|+\sum_{j=1}^l|U^{(n)}_j|\big). \end{aligned}\tag{3.29}$$

Taking into account (2.9), (2.10), (3.15) and (3.20) we obtain that

$$\begin{aligned}|\Phi_l(0,...,0)|\leq\max\big(&LT+|F_0(0)|+|G_0(0)|,\,|\Phi_{l+1}(0,...,0)|\\ &+L(T/n)^{1/2}(N+M)E|X_1|\big)\end{aligned}$$

with $|\Phi_{\ell(n)}(x_1,...,x_{\ell(n)})|\leq LT+|F_0(0)|$, and so for each $l=0,1,...,\ell(n)$,

$$|\Phi_l(0,...,0)|\leq LT+|F_0(0)|+|G_0(0)|+L(Tn)^{1/2}E|X_1|. \tag{3.30}$$

Now, collecting (3.22)–(3.26) and (3.28)–(3.30) we arrive at (3.18) for $k=l$ completing the induction step and proving (3.18) for all $k=0,1,...,\ell(n)$. The estimate (3.19) follows from (3.22)–(3.25) and (3.27)–(3.30) in the same way.

The proof of the equality (3.11) is similar but simpler taking into account that $\tilde U^{(n)}_j=B(t_{j+1}(n))-B(t_j(n))$, $j=k+1,...,\ell(n)-1$ are independent both of $\mathcal{F}^B_{t_k(n)}$ and of $\mathcal{G}^B_k$, and so we can construct in the same way as in Lemma 3.1 of [10] some functions $\tilde\Phi_k(x_1,...,x_k)$, $k=1,2,...,\ell(n)$ and a constant $\tilde\Phi_0$ such that for $k=0,1,...,\ell(n)$,

$$V^B_{\mathcal{S},k}=V^B_{n,k}=\tilde\Phi_k(\tilde U^{(n)}_1,...,\tilde U^{(n)}_k)$$

where $V^B_{\mathcal{S},0}=V^B_{\mathcal{S}}$, $V^B_{n,0}=V^B_n$ and $V^B_{\mathcal{S},k}$, $V^B_{n,k}$ are obtained via the dynamical programming recursive formulas similar to (3.12)–(3.14). $\square$

3.3. **Remark.** The functions Φ_l and the constant Φ_0 constructed via the recursion (3.20) is the price process and the value of Dynkin's game with the payoff R^U_n given by (3.7) but with respect to the probability which makes $U^{(n)}_1,...,U^{(n)}_{\ell(n)}$ independent identically distributed (i.i.d) sequence where each $U^{(n)}_j$ has the same distribution as $U^{(n)}_1=\sum_{j=1}^{M+N}X_j$. Hence, (3.18) and (3.19) estimate the error which originates from replacement of ψ-mixing by independency of large blocks.

4. Characteristic functions and strong approximations

Next, we will need certain estimates for characteristic functions of normalized sums of random vectors X_k, $k\geq1$ which play a crucial role in strong approximations via the approach of [1] and [3]. We remark that these estimates hold true with the α-mixing coefficient which does not exceed quarter of the ψ-mixing coefficient (see [4], Section 1.1) but we use the latter in the results below in order to avoid additional definitions and since, anyway, the estimates of Section 3 rely on ψ-mixing. Note that in this section the symbol i will be reserved for $\sqrt{-1}$. For $k\leq l\leq\ell(n)$ and $y\in\mathbb{R}^d$ set

$$\varphi^{(n)}_{kl}(y)=E\exp\big(i(T/n)^{1/2}\langle y,\sum_{j=k}^l U^{(n)}_j\rangle\big)$$

and

$$f^{(n)}_{kl}(y)=\exp\big(-\frac{1}{2}(t_{l+1}(n)-t_k(n))\langle y,Ay\rangle\big)$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product and, recall, $t_j(n)=jT(M+N)/n$. The following result is proved as Theorem 3.24 in [3] (see also [9], Proposition 4.3) and we refer the reader for details there.

4.1. **Lemma.** *Let $X_1,X_2,...,EX_1=0$ be a ψ-mixing stationary sequence of d-dimensional random vectors such that for some $\kappa_1,\kappa_2,K_1,K_2>0$, $\kappa_1\le 1,\kappa_2<1/4$ and all $n\in\mathbb{N}$,*

$$E|X_1|^{2+\kappa_1}\le K_1\quad\textit{and}\quad\psi(n)\le K_2n^{-(1+\kappa_2)(1+2/\kappa_1)}. \tag{4.1}$$

Then the two series in (2.4) converge absolutely and there exists $C=C_T>0$ such that for all $k\le l\le n$,

$$|\varphi^{(n)}_{kl}(y)-f^{(n)}_{kl}(y)|\le C|y|q_T\big(|y|,(M+N)(k-l)\big) \tag{4.2}$$

where

$$q_T^{\kappa_1,\kappa_2}(b,m)=K_1(m/n)^{9/8}b^{5/4}d^{9/4}\big((m/n)^{9/8}(bd)^{9/4}m^{-3/16}+m^{-\frac{1}{56}\kappa_1\kappa_2}\big)$$

and under the condition (2.13) of Theorem 2.1 above we can take here $\kappa_1=1$ and $\kappa_2=1/5$.

It follows from [5] that for $k=1,...,\ell(n)$,

$$E\big|E\big(\exp(i(T/n)^{1/2}\langle y,Y^{(n)}_k\rangle)\big|\mathcal{F}^U_{k-1}\big)-E\exp(i(T/n)^{1/2}\langle y,Y^{(n)}_k\rangle)\big|\le 2\pi\psi(M) \tag{4.3}$$

which can be proved with 8 in place of 2π applying the simple real case proof from [12] (where 4 stands in place of 2π) to the real and the imaginary parts of the left hand side of (4.3). Next,

$$E\big|\exp(i(T/n)^{1/2}\langle y,U^{(n)}_k\rangle)-\exp(i(T/n)^{1/2}\langle y,Y^{(n)}_k\rangle)\big|\le$$
$$E\big|\exp(i(T/n)^{1/2}\langle y,Z^{(n)}_k\rangle)-1\big|\le(T/n)^{1/2}|y|E|Z^{(n)}_k|\le(T/n)^{1/2}M|y|E|X_1|$$

which together with (4.2) and (4.3) yield

$$E\big|E\big(\exp(i(T/n)^{1/2}\langle y,U^{(n)}_k\rangle)\big|\mathcal{F}^U_{k-1}\big)-f^{(n)}_{kk}(y)\big|\le\rho(|y|,M,M+N,n) \tag{4.4}$$

where

$$\rho(b,m,l,n)=Cbq_T(b,l)+2\pi\psi(m)+2(T/n)^{1/2}mbE|X_1|.$$

Next, let $B_A(t)$, $t\in[0,T]$ be the d-dimensional Brownian motion with the covariance matrix A. For $k=1,2,...,\ell(n)$ set $\tilde W^{(n)}_k=B_A(t_k(n))-B_A(t_{k-1}(n))$. Then by direct estimates it follows that for any $D>0$,

$$P\{\tilde W^{(n)}_k\ge D/4\}\le e^{-C_TD^2\ell(n)} \tag{4.5}$$

for some $C_T>0$ depending only on T. It follows from (4.4) and (4.5) together with Theorem 4.6 of [3] (which is an advanced version of Theorem 1 of [1]) that we can redefine the sequence $\{U^{(n)}_k,\,k=1,...,\ell(n)\}$ on a richer probability space where there exists a sequence $\{W^{(n)}_k,\,k=1,...,\ell(n)\}$ of independent random vectors having the same distribution as $\tilde W^{(n)}_k$ above and such that

$$P\big\{\big|(T/n)^{1/2}U^{(n)}_k-W^{(n)}_k\big|\ge\beta\big\}\le\beta,\ k=1,...,\ell(n) \tag{4.6}$$

where

$$\beta=10dD^{-1}\log D+4\big(\rho(D,M,M+N,n)\big)^{1/2}D^d+e^{-C_TD^2\ell(n)} \tag{4.7}$$

with any $D\geq 10^8d$ (where the estimate holds true from $k=1$ and not from $k=2$ as written in [3]). Moreover, on the same probability space there exists also a sequence $\xi_1,\xi_2,...$ of independent uniformly distributed on $[0,1]$ random variables independent of the sequence $U_1^{(n)},...,U_{\ell(n)}^{(n)}$ such that $W_k^{(n)}$ is measurable with respect to the σ-algebra $\mathcal{H}_{kk}$ generated by $U_k^{(n)}$ and ξ_k, $k=1,2,...$ and $W_k^{(n)}$ is independent of the σ-algebra $\mathcal{H}_{1,k-1}$ generated by $U_1^{(n)},...,U_{k-1}^{(n)}$ and $\xi_1,...,\xi_{k-1}$, $k=2,3,....$

From (4.6) it is easy to derive (see (4.23) in [9]) that for any $R>0$ and $k=1,2,...,\ell(n)$,

$$\begin{aligned}&E\big|(T/n)^{1/2}U_k^{(n)}-W_k^{(n)}\big|^2 \\ &\leq\lambda\overset{\text{def}}{=}\beta^2+8R^2\beta+4R^{-1}\big((T/n)^{3/2}E|U_k^{(n)}|^3+E|W_k^{(n)}|^3\big).\end{aligned} \tag{4.8}$$

It follows from the Kolmogorov existence (extension) theorem (see, for instance, [2]) that without changing distributions we can redefine the sequence $X_1,X_2,...$ and $\xi_1,\xi_2,...$ on a richer probability space where there exists a d-dimensional Brownian motion B_A, $t\geq 0$ with the covariance matrix A such that the joint distribution of

$$U_k^{(n)}=\sum_{i=1}^{M+N}X_{i+(k-1)(M+N)},\ \xi_k,\ W_k^{(n)}=B_A(t_k(n))-B_A(t_{k-1}(n)),\ k=1,2,...$$

is the same as before, in particular, (4.8) holds true and, again, $W_k^{(n)}$ is measurable with respect to the σ-algebra $\mathcal{H}_{kk}$ and it is independent of $\mathcal{H}_{1,k-1}$ using the same notations as before. Indeed, according to the Kolmogorov theorem we can construct such triple $\big(X_1,X_2,...;\,\xi_1,\xi_2,...;\,B_A(t),t\geq 0\big)$ provided we impose consistent conditions on joint distributions which holds true here since we already have a triple $(U_k^{(n)},\,\xi_k,\,W_k^{(n)},\,k\geq 1)$ satisfying required conditions.

Let $\mathcal{T}_n^{\mathcal{H}}$ be the space of stopping times with respect to the filtration $\big\{\mathcal{H}_{1,0}=\{\emptyset,\Omega\},\,\mathcal{H}_{1,k},\,k=1,...,\ell(n)\big\}$ taking values $0,t_1(n),t_2(n),...,t_{\ell(n)}(n)$. Set

$$V^U_{\mathcal{T}_n^{\mathcal{H}}}=\inf_{\zeta\in\mathcal{T}_n^{\mathcal{H}}}\sup_{\eta\in\mathcal{T}_n^{\mathcal{H}}}ER_n^U(\zeta,\eta)\quad\text{and}\quad V^B_{\mathcal{T}_n^{\mathcal{H}}}=\inf_{\zeta\in\mathcal{T}_n^{\mathcal{H}}}\sup_{\eta\in\mathcal{T}_n^{\mathcal{H}}}ER^B(\zeta,\eta) \tag{4.9}$$

where R_n^U and R^B are given by (3.7) and (2.7), respectively.

4.2. **Lemma.** *For any N,M,n and T as above,*

$$V^U_{\mathcal{T}_n^{\mathcal{H}}}=V_n^U\quad\textit{and}\quad V^B_{\mathcal{T}_n^{\mathcal{H}}}=V_n^B \tag{4.10}$$

where V_n^U and V_n^B are the same as in Lemma 3.2.

Proof. As in Lemma 3.2 we obtain first $V^U_{\mathcal{T}_n^{\mathcal{H}}}$ and $V^B_{\mathcal{T}_n^{\mathcal{H}}}$ by means of the dynamical programming procedure, so that $V^U_{\mathcal{T}_n^{\mathcal{H}}}=V^U_{\mathcal{T}_n^{\mathcal{H}},0}$ and $V^B_{\mathcal{T}_n^{\mathcal{H}}}=V^B_{\mathcal{T}_n^{\mathcal{H}},0}$ where

$$V^U_{\mathcal{T}_n^{\mathcal{H}},\ell(n)}=F_{t_{\ell(n)}(n)}(S_n^U),\ V^B_{\mathcal{T}_n^{\mathcal{H}},\ell(n)}=F_{t_{\ell(n)}(n)}(B_A) \tag{4.11}$$

and recursively

$$V^U_{\mathcal{T}_n^{\mathcal{H}},k}=\min\Big(G_{t_k(n)}(S_n^U),\max\big(F_{t_k(n)}(S_n^U),E\big(V^U_{\mathcal{T}_n^{\mathcal{H}},k+1}\big|\mathcal{H}_{1,k})\big)\Big) \tag{4.12}$$

and

$$V^B_{\mathcal{I}^{\mathcal{H}}_n,k} = \min\Big(G_{t_k(n)}(B_A), \max\big(F_{t_k(n)}(B_A), E\big(V^B_{\mathcal{I}^{\mathcal{H}}_n,k+1}\big|\mathcal{H}_{1,k}\big)\big)\Big). \tag{4.13}$$

Next, we prove by the backward induction that there exist integrable random variables Ψ^U_k, Ψ^B_k and constants Ψ^U_0, Ψ^B_0 such that Ψ^U_k and Ψ^B_k are measurable with respect to $\mathcal{F}^U_k$ and $\mathcal{G}^B_k$, respectively,

$$V^U_{\mathcal{I}^{\mathcal{H}}_n,k} = \Psi^U_k = V^U_{n,k} \tag{4.14}$$

and

$$V^B_{\mathcal{I}^{\mathcal{H}}_n,k} = \Psi^B_k = V^B_{n,k} \tag{4.15}$$

where $V^U_{n,k}$ is given by (3.14) and $V^B_{n,k}$ is obtained by similar recursive formulas. For $k = \ell(n)$ we obtain (4.14) and (4.15) from (4.11) setting

$$\Psi^U_{\ell(n)} = F_{t_{\ell(n)}(n)}(S^U_n) \quad \text{and} \quad \Psi^B_{\ell(n)} = F_{t_{\ell(n)}(n)}(B_A).$$

Suppose that some integrable random variables Ψ^U_k and Ψ^B_k are already constructed for all $k \geq l+1$ so that (4.14) and (4.15) hold true for such k. Set

$$\Psi^U_l = \min\Big(G_{t_l(n)}(S^U_n), \max\big(F_{t_l(n)}(S^U_n), E\big(\Psi^U_{l+1}\big|\mathcal{F}^U_l\big)\big)\Big) \tag{4.16}$$

and

$$\Psi^B_l = \min\Big(G_{t_l(n)}(B_A), \max\big(F_{t_l(n)}(B_A), E\big(\Psi^B_{l+1}\big|\mathcal{G}^B_l\big)\big)\Big). \tag{4.17}$$

Observe, that by (2.9), (2.10) and (2.13) the random variables Ψ^U_l and Ψ^B_l are integrable. By (3.14) and the corresponding formula for $V^B_{n,l}$ we have that $V^U_{n,l} = \Psi^U_l$ and $V^B_{n,l} = \Psi^B_l$. Hence, in order to obtain (4.14) and (4.15) for $k = l$ it suffices to show that with probability one

$$E(\Psi^U_{l+1}|\mathcal{F}^U_l) = E(\Psi^U_{l+1}|\mathcal{H}_{1,l}) \tag{4.18}$$

and

$$E(\Psi^B_{l+1}|\mathcal{G}^B_l) = E(\Psi^B_{l+1}|\mathcal{H}_{1,l}). \tag{4.19}$$

In order to prove (4.18) observe that $\mathcal{H}_{1,l}$ is generated by $\mathcal{F}^U_l$ and the σ-algebra $\mathcal{K}_l$ which is generated by $\xi_1, ..., \xi_l$. Since $\mathcal{H}_{1,l}$ is also generated by disjoint unions $\cup_i A_i \cap B_i$ where $A_i \in \mathcal{F}^U_l$ and $B_i \in \mathcal{K}_l$ and since Ψ^U_{l+1} can be approximated by linear combinations of indicators of disjoint events from the σ-algebra $\mathcal{Q}_{l+1}$ generated by Ψ_{l+1} it follows that (4.18) will hold true if for any $A \in \mathcal{F}^U_l$, $B \in \mathcal{K}_l$ and $Q \in \mathcal{Q}_{l+1}$,

$$\int_{A\cap B} P(Q|\mathcal{F}^U_l)dP = \int_{A\cap B} P(Q|\mathcal{H}_{1,l})dP. \tag{4.20}$$

Recall, that by Theorem 4.6 of [3], $\mathcal{K}_l$ is independent of $\mathcal{F}^U_{l+1}$ which contains both $\mathcal{F}^U_l$ and $\mathcal{Q}_{l+1}$, and so B is independent both of $A \cap Q$ and of the σ-algebra $\mathcal{F}^U_l$. Hence,

$$\begin{aligned}\textstyle\int_{A\cap B} P(Q|\mathcal{F}^U_l)dP = \int_A E\big(P(Q|\mathcal{F}^U_l)\mathbb{I}_B\big|\mathcal{F}^U_l\big)dP = P(B)\int_A P(Q|\mathcal{F}^U_l)dP \\ = P(B)P(A\cap Q) = P(B\cap A\cap Q) = \textstyle\int_{A\cap B} P(Q|\mathcal{H}_{1,l})dP\end{aligned}$$

obtaining (4.20).

In order to prove (4.19), recall, that by the induction hypotesis Ψ^B_{l+1} is measurable with respect to the σ-algebra $\mathcal{G}^B_{l+1}$ which means that $\Psi^B_{l+1}=Q(W_1^{(n)},...,W_{l+1}^{(n)})$ for some Borel function $Q(x_1,...,x_{l+1})$. Set $\hat{Q}(x_1,...,x_l)=EQ(x_1,...,x_l,W_{l+1}^{(n)})$. Since $W_{l+1}^{(n)}$ is independent both of $\mathcal{G}^B_l$ and of $\mathcal{H}_{1,l}$ then by the easy well known part of Lemma 3.1 both the left and the right hand side of (4.19) equals $\hat{Q}(W_1^{(n)},...,W_l^{(n)})$, completing the induction step and the proof of Lemma 4.2. $\square$

4.3. **Remark.** Showing (4.18) we proved the following general fact which should be listed among major properties of conditional expectations but we could not find it in standard textbooks. Namely, let a σ-algebra $\mathcal{H}$ be generated by a σ-algebra $\mathcal{F}$ and by another σ-algebra $\mathcal{K}$ which is independent of the σ-algebra generated by $\mathcal{F}$ and by the σ-algebra generated by a random variable Ψ with $E|\Psi|<\infty$. Then

$$E(\Psi|\mathcal{F})=E(\Psi|\mathcal{H}).$$

Next, we bridge between values of Dynkin's games with payoffs based on normalized sums S_n^U and games with payoffs based on the process B_A which is the only place where we use the strong approximation theorem from [3] described above. The crucial point here is that now we can deal with games where stopping times are defined with respect to the same filtration $\{\mathcal{H}_{1,k},\, k\geq 1\}$.

4.4. **Lemma.** *For any M,N and n as above,*

$$\big|V^U_{\mathcal{T}^{\mathcal{H}}}-V^B_{\mathcal{T}^{\mathcal{H}}}\big|\leq L\ell(n)\sqrt{\lambda} \tag{4.21}$$

where λ is the same as in (4.8) (and it depends also on T in view of (4.7) and (4.8)).

Proof. By (2.7), (2.9), (3.7), (4.8), (4.9) and the Cauchy-Schwarz inequality,

$$\begin{aligned}\big|V^U_{\mathcal{T}^{\mathcal{H}}}-V^B_{\mathcal{T}^{\mathcal{H}}}\big| &\leq \sup_{\zeta\in\mathcal{T}_n^{\mathcal{H}}}\sup_{\eta\in\mathcal{T}_n^{\mathcal{H}}}E|R_n^U(\zeta,\eta)-R^B(\zeta,\eta)|\\ &\leq LE\max_{1\leq k\leq\ell(n)}|S_n^U(t_k(n))-B_A(t_k(n))|\\ &\leq L\sum_{k=1}^{\ell(n)}E|(T/n)^{1/2}U_k^{(n)}-W_k^{(n)}|\leq L\ell(n)\sqrt{\lambda}\end{aligned}$$

proving (4.21). $\square$

5. Remaining estimates

In this section we compare $V^X_{\mathcal{S}}$ with V^X_n and $V^B_{\mathcal{S}}$ with V^B and collect all estimates into the one which after some optimisation gives the right hand side of (2.14) and completes the proof of Theorem 2.1 under the conditions (2.9) and (2.10). After that we show how to work under the conditions (2.11), (2.12) and (2.15) in order to obtain, essentially, the same estimate and, finally, we deal with the case of independent random vectors.

5.1. **Lemma.** *There exists $C=C_T>0$ such that if M,N and n are as before then*

$$|V^X_{\mathcal{S}}-V^X_n|\leq C\big(\ell(n)\big)^{-1/6} \tag{5.1}$$

and

$$|V^B_{\mathcal{S}}-V^B|\leq C\big(\ell(n)\big)^{-1/6}. \tag{5.2}$$

Proof. For any $\xi \in \mathcal{T}_n^X$ set

$$\xi^{\mathcal{S}} = \min\{kT(M+N)/n : \, k \geq \frac{n\xi}{T(M+N)}\}$$

which defines a stopping time from $\mathcal{S}_n^X$ satisfying

$$\xi + \frac{T}{n} \geq \xi^{\mathcal{S}} \geq \xi. \tag{5.3}$$

Next, we proceed similarly to the proof of Lemma 3.6 in [10]. Namely, since $\mathcal{S}_n^X \subset \mathcal{T}_n^X$ we obtain from (2.6) that

$$V_n^X \geq \inf_{\zeta \in \mathcal{T}_n^X} \sup_{\eta \in \mathcal{S}_n^X} ER_n^X(\zeta, \eta). \tag{5.4}$$

Then for any $\delta > 0$ there exists $\zeta_\delta \in \mathcal{T}_n^X$ such that

$$V_n^X \geq \sup_{\eta \in \mathcal{S}_n^X} ER_n^X(\zeta_\delta, \eta) - \delta, \tag{5.5}$$

and so by (3.8),

$$\begin{aligned} V_n^X \geq \sup\nolimits_{\eta \in \mathcal{S}_n^X} ER_n^X(\zeta_\delta^{\mathcal{S}}, \eta) - \delta - \sup\nolimits_{\eta \in \mathcal{S}_n^X} E\big(R_n^X(\zeta_\delta^{\mathcal{S}}, \eta) \\ -R_n^X(\zeta_\delta, \eta)\big) \geq V_{\mathcal{S}}^X - \delta - \sup\nolimits_{\eta \in \mathcal{S}_n^X} J_1(\zeta_\delta, \eta) \end{aligned} \tag{5.6}$$

where for any $\zeta \in \mathcal{T}_n^X$ and $\eta \in \mathcal{S}_n^X$,

$$J_1(\zeta, \eta) = E\big(R_n^X(\zeta^{\mathcal{S}}, \eta) - R_n^X(\zeta, \eta)\big).$$

By (5.3),

$$R_n^X(\zeta, \eta) = G_\zeta(S_n^X) \quad \text{whenever} \quad R_n^X(\zeta^{\mathcal{S}}, \eta) = G_{\zeta^{\mathcal{S}}}(S_n^X).$$

Hence, by (2.3), (2.10) and (5.3),

$$\begin{aligned} R_n^X(\zeta^{\mathcal{S}}, \eta) - R_n^X(\zeta, \eta) \leq |G_{\zeta^{\mathcal{S}}}(S_n^X) - G_\zeta(S_n^X)| \\ \leq LTn^{-1}\big(1 + (T/n)^{1/2} \textstyle\sum_{j=1}^n |X_j|\big) + L(T/n)^{1/2} \max_{1 \leq k \leq \ell(n)} I_k \end{aligned} \tag{5.7}$$

where for all $a, b \in \mathbb{N}$, $b \geq a$,

$$I_k = \max_{0 \leq l \leq M+N} |\Delta\big(1+(k-1)(M+N), l+(k-1)(M+N)\big)| \quad \text{and} \quad \Delta(a,b) = \sum_{j=a}^b X_j.$$

It follows from [23] (see also Theorem 2 and Remark 2 in Section 1.4.1 of [4]) that

$$E|\Delta(a,b)|^3 \leq C(b-a)^{3/2} \tag{5.8}$$

for some $C > 0$ provided $b > a$. Hence, by the maximal inequality of Theorem 6 from Section 1.4.3 in [4] for any $m > 1$ and $k = 1, ..., \ell(n)$,

$$EI_k^3 \leq \tilde{C}(M+N)^{3/2} \tag{5.9}$$

with some $\tilde{C} > 0$. This together with the Hölder inequality and stationarity of the sequence X_j, $j \geq 1$,

$$\begin{aligned} E \max\nolimits_{1 \leq k \leq \ell(n)} I_k \leq \big(E \max\nolimits_{1 \leq k \leq \ell(n)} I_k^3\big)^{1/3} \\ \leq \big(\ell(n) E I_1^3\big)^{1/3} \leq \tilde{C}^{1/3}(\ell(n))^{1/3}(M+N)^{1/2}. \end{aligned} \tag{5.10}$$

Since $\delta > 0$ in (5.6) is arbitrary we derive from (5.6), (5.7) and (5.10) that

$$V_n^X \geq V_{\mathcal{S}}^X - LT(n^{-1} + (T/n)^{1/2} E|X_1|) - LT^{1/2}\tilde{C}^{1/3}(\ell(n))^{-1/6}. \tag{5.11}$$

By [17], we can represent V_n^X not only in the form (2.6), but also as

$$V_n^X = \sup_{\eta\in\mathcal{T}_n^X}\inf_{\zeta\in\mathcal{T}_n^X} ER_n^X(\zeta,\eta) \le \inf_{\zeta\in\mathcal{S}_n^X} ER_n^X(\zeta,\eta_\delta)+\delta$$

for each $\delta > 0$ and some $\eta_\delta \in \mathcal{T}_n^X$. Introducing $\eta_\delta^{\mathcal{S}}$ and employing the same arguments as above we obtain an inequality in the other than (5.11) direction with minuses replaced by pluses there and (5.1) follows.

The proof of (5.2) is the same as above but in place of the moment and maximal inequalities for mixing sequences from [4] which we employed in order to derive (5.8)–(5.10) we can take advantage now of the corresponding martingale (maximal) moment inequalities for increments of the Brownian motion B_A (see, for instance, [7]). □

Now, writing

$$\begin{aligned}(5.12)\qquad |V_n^X - V^B| \le |V_n^X - V_{\mathcal{S}}^X| + |V_{\mathcal{S}}^X - V_n^U| + |V_n^U - V^U_{\mathcal{T}_n^{\mathcal{H}}}| \\ +|V^U_{\mathcal{T}_n^{\mathcal{H}}} - V^B_{\mathcal{T}_n^{\mathcal{H}}}| + |V^B_{\mathcal{T}_n^{\mathcal{H}}} - V_n^B| + |V_n^B - V_{\mathcal{S}}^B| + |V_{\mathcal{S}}^B - V^B|\end{aligned}$$

and collecting (5.1), (3.10), (4.10), (4.21), (4.10), (3.11) and (5.2), respectively, we obtain that

$$(5.13)\qquad |V_n^X - V^B| \le C\gamma(M,N,n,R,D)$$

for some $C = C_T > 0$ where

$$(5.14)\quad \gamma(M,N,n,R,D) = (n/N)^{-1/6} + \psi(M)n^{3/2}N^{-1} + MN^{-2}n^{3/2} + nN^{-1}\sqrt{\lambda}$$

where $\lambda = \lambda(M,N,n,R,D)$ is defined in (4.8) and (as it has to be chosen anyway) $N > \sqrt{n} > M$. Observe that

$$E|W_k^{(n)}|^{2m} = \big(T(M+N)/n\big)^m E|B_A(1)|^{2m}$$

and by (5.8),

$$E|U_k^{(n)}|^3 \le C(M+N)^{3/2}.$$

Hence, we can estimate λ in (4.8) by

$$(5.15)\qquad \lambda \le \beta^2 + 8R^2\beta + 4R^{-1}C_T(N/n)^{3/2}$$

for some $C_T > 0$. Since β is small and R will be large we can disregard β^2 in (5.15) and only take care about other terms. Since D in the definition (4.7) of β will be chosen of order of some positive power of n, the exponential term there is very small and can be disregarded. Using (2.13), (5.13)–(5.15) and the trivial inequality

$$\Big(\sum_{1\le j\le k} a_j\Big)^{1/l} \le \sum_{1\le j\le k} a_j^{1/l} \quad \text{for any} \quad a_1,...,a_k \ge 0,\ l\ge 1$$

we obtain after the above simplifications that

$$\gamma(M,N,n,R,D) \le \tilde{C}\sqrt{\log D}\,\tilde{\gamma}(M,N,n,R,D)$$

where

$$\begin{aligned}(5.16)\qquad \tilde{\gamma}(M,N,n,R,D) = (n/N)^{-1/6} + M^{-4}N^{-1}n^{3/2} + MN^{-2}n^{3/2} \\ +R^{-1/2}(n/N)^{1/4} + nN^{-1}RD^{-1/2} + RD^{\frac{1}{8}(4d+9)}N^{-\frac{31}{64}}n^{7/16} \\ +RD^{\frac{1}{16}(8d+9)}N^{-\frac{403}{560}}n^{\frac{23}{32}} + nN^{-1}M^{-1}RD^{d/2} + N^{-1}n^{7/8}M^{1/4}RD^{\frac{1}{4}(2d+1)}.\end{aligned}$$

Equating first, fourth, fifth, and seventh terms in the right hand side of (5.16) we obtain that we should set

$$N \sim n^{1-\frac{1}{2240d+4446}},\ R \sim n^{\frac{5}{13440d+26556}} \text{ and } D \sim n^{\frac{2}{1120d+2223}}.$$

The remaining terms in (5.16) will be of smaller order than the above ones if we set, for instance, $M \sim n^{3/8}$ which yields the estimate of $\tilde{\gamma}$ of order $n^{-1/(13440d+26556)}$ implying (2.14).

Next, we discuss the proof under the conditions (2.11) and (2.12) in place of (2.9) and (2.10). Observe that we needed these conditions in order to obtain the estimate (3.10) of Lemma 3.2 (see (3.15)–(3.17), (3.24), (3.25), (3.29) and (3.30)), in order to derive (4.21) in Lemma 4.4 and in (5.7) which was needed for estimates (5.1) and (5.2) of Lemma 5.1. All expressions which we have to estimate now have the form

$$A_1 = E(|\Theta| + \mathbb{I}_{|\Theta|>1})e^{|\Upsilon|}$$

for some random vectors Θ and Υ, so that we already obtained and used (or can obtain in the same way as we did already above) estimates of $E|\Theta|$ and $E|\Theta|^2$. By the Cauchy–Schwarz and Chebyshev inequalities

$$A_1 \le \big((E|\Theta|^2)^{1/2} + (P\{|\Theta|>1\})^{1/2}\big)(Ee^{2|\Upsilon|})^{1/2} \le 2(E|\Theta|^2)^{1/2}(Ee^{2|\Upsilon|})^{1/2}.$$

So it remains only to estimate the expectation of the exponential term. The expectations of exponentials we have to deal with are of two types

$$A_2 = E \max_{1\le k\le n} \exp\big(K(T/n)^{1/2}|\sum_{j=1}^{k} X_j|\big) \text{ and } A_3 = E \sup_{0\le t\le T} \exp\big(K|B_A(t)|\big) \tag{5.17}$$

where $K>0$ is a constant. It is well known that the second expectation in (5.17) is finite and its estimate can be obtained writing, first, $K|B_A(t)| \le \tilde{K}\sum_{j=1}^{d}|B_j(t)|$ for some $\tilde{K}>0$ where $B_1(t),...,B_d(t)$ are independent one dimensional standard Brownian motions, so that

$$E \sup_{0\le t\le T} \exp(K|B_A(t)|) \le \big(E \sup_{0\le t\le T} \exp(\tilde{K}|B_1(t)|\big)^d \le \big(2E \sup_{0\le t\le T} \exp(\tilde{K}B_1(t))\big)^d$$

and the latter expression can be easily estimated using the exponential martingale (see, for instance, Section 4 in [10]). In fact, such expression is finite not only for B_A but for any solution of a stochastic differential equation with smooth bounded coefficients.

In order to estimate A_2 we write using the Cauchy–Schwarz inequality

$$\begin{aligned} A_2 &\le \log n + E\textstyle\sum_{k=1}^{n} \mathbb{I}_{\Gamma_k(\log n)} \exp\big(K(T/n)^{1/2}|\sum_{j=1}^{k} X_j|\big) \\ &\le \log n + \textstyle\sum_{k=1}^{n}\big(P(\Gamma_k(\log n))\big)^{1/2}\big(E\exp\big(2K(T/n)^{1/2}|\sum_{j=1}^{k} X_j|\big)\big)^{1/2} \end{aligned} \tag{5.18}$$

where $\Gamma_k(x) = \{K(T/n)^{1/2}|\sum_{j=1}^{k} X_j| \ge x\}$. Next, we use the boundedness assumption (2.15) together with moment inequalities from Section 1.4.1 in [4] (see also Theorem 3.20 in [3]) which enable us to employ Theorem 5 from Section 1.4.2 of [4] yielding that for some $a,b>0$ depending on T but not on $k\le n$ and n,

$$P(\Gamma_k(x)) \le ae^{-bx^2}$$

for any $x\ge 0$. This together with (5.18) gives that

$$A_2 \le \log n + C$$

where $C > 0$ does not depend on n. Hence, all estimates we are dealt with in the way of proving (2.14) under (2.9) and (2.10) remain true also under (2.11) and (2.12)with may be an extra factor $(\log n)^{1/2}$, i.e. (2.14) remains valid with this extra factor under (2.11) and (2.12) in place of (2.9) and (2.10) provided we assume also (2.15).

If $X_1, X_2, ...$ are independent then $\psi(m) \equiv 0$, we can take $M = 0$ and the estimate (4.2) of Lemma 4.1 can be improved leading to a better estimate in (2.14). Namely, by independency and stationarity,

$$\varphi_{kl}^{(n)} = \left(E \exp\big(i(T/n)^{1/2}\langle y, X_1\rangle\big)\right)^{N(k-l)}$$

and assuming the nondegeneracy of the covariance matrix of X_1 we obtain via standard estimates (see Lemma 5.1 in Ch.5 of [19] or §6, Ch.III in [21]) that for any $1 \le k \le l \le n$,

$$|\varphi_{kl}^{(n)}(y) - f_{kl}^{(n)}(y)| \le C\big(|y|^3 n^{-3/2} + n^{-1/2}(\log n)^{3/2}\big) \tag{5.19}$$

for some $C > 0$ independent of y, k, l and n. Then, in place of $\tilde{\gamma}$ given by (5.16) we have to estimate

$$\begin{aligned}\tilde{\gamma}(M, N, n, D, R) = (\log n)^{3/8}\big((n/N)^{-1/6} + R^{-1/2}(n/N)^{1/4} \\ + N^{-1}nRD^{-1/2} + N^{-1}n^{5/8}RD^{\frac{1}{4}(2d+3)} + N^{-1}n^{7/8}RD^{d/2}\big).\end{aligned} \tag{5.20}$$

Equating first, second, third and fifth terms in the right hand side of (5.20) we obtain that we should set

$$N \sim n^{1-\frac{1}{16d+40}},\ R \sim n^{\frac{5}{32d+80}},\ \text{and } D \sim n^{\frac{3}{4d+10}}.$$

The remaining term in (5.20) will be of smaller order than the above ones which yields the estimate of $\tilde{\gamma}$ of order $n^{-1/96(d+1)}$, completing the proof of Theorem 2.1. □

References

[1] I.Berkes and W.Philipp, *Approximation theorems for independent and weakly dependent random vectors,* Ann. Probab. **7** (1979), 29–54.

[2] P.Billingsley, *Probability and Measure,* J.Willey, New York, 1986.

[3] H.Dehling and W.Philipp, *Empirical process technique for dependent data,* in:*Empirical Process Technique for Dependent Data* (ed. H.G.Dehling, T.Mikosch, M.Sorenson), p.p. 3–113, Birkhäuser, Boston, 2002.

[4] P.Doukhan, *Mixing. Properties and Examples*, Springer-Verlag, New York, 1994.

[5] A.Dvoretzky, *Asymptotic normality for sums of dependent random variables,* in: Proc. Sixth Berekeley Symp. Math. Stat. Probab., vol.II (1970), 513–535.

[6] L.Heinrich, *Mixing properties and central limit theorem for a class of nonidentical piece-wise monotonic C^2-transformations*, Math. Nachricht. **181** (1996), 185–214.

[7] N.Ikeda and S.Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland, Amsterdam, 1989.

[8] Yu.Kifer, *Game options,* Finance and Stoch. **4** (2000), 443–463.

[9] Yu.Kifer, *L^2 diffusion approximation for slow motion in averaging,* Stoch. Dynam. **3** (2003), 213–246.

[10] Yu.Kifer, *Error estimates for binomial approximations of game options,* Ann. Appl. Probab. (2006)

[11] J.Komlós, P.Major and G.Tusnády, *An approximation of partial sums of independent RV's and the sample DF. II,* Z. Warsch. verw. Geb. **34** (1976), 33–56.

[12] J.Kuelbs and W.Philipp, *Almost sure invariance principles for partial sums of mixing B-valued random variables,* Ann. Probab. **8** (1980), 1003–1036.

[13] H.Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations,* Acad. Press, 1977.

[14] D.Lamberton and G.Pagès, *Sur l'approximation of réduites*, Ann. l'Inst. H.Poincaré **26** (1990), 331–355.

[15] J.P.Lepeltier and J.P.Maingueneau, *Le jeu de Dynkin en theorie generale sans l'hypothese de Mokobodski,* Stochastics **13** (1984), 24–44.

[16] S.Mulinacci and M.Pratelli, *Functional convergence of Snell envelopes: applications to American options approximations*, Finance and Stoch. **2**(1998), 311–327.

[17] Y.Ohtsubo, *Optimal stopping in sequential games with or without a constraint of always terminating,* Math. Oper. Res. **11** (1986), 591–607.

[18] M.Peligrad, *Invariance principles for mixing sequences of random variables*, Ann. Probab. **10** (1982), 968–981.

[19] V.V.Petrov, *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*, Oxford Univ. Press, Oxford, 1995.

[20] W.Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

[21] A.N.Shiryaev, *Probability*, Springer-Verlag, New York, 1984.

[22] V.Strassen, *Almost sure behavior of sums of independent random variables and martingales*, Proc. Fifth Berkeley Symp. Math. Stat. Probab. vol.II, Part 1, p.p. 315–343 (1967).

[23] R.Yokoyama, *Moment bounds for stationary mixing sequences*, Z. Wahrsch. Verw. Gebiete **52** (1980), 45–57.

[24] A.Yu.Zaitsev, *Multidimensional version of a result of Sakhanenko in the invariance principle for vectors with finite exponential moments, I–III,* Theory Probab. Appl. **45** (2000), 624–641; 45 (2001), 490–514, 676–698.

Institute of Mathematics, The Hebrew University, Jerusalem 91904, Israel
E-mail address: kifer@math.huji.ac.il