*Journal of Theoretical Probability, Vol. 9, No. 4, 1996*

# On the Interrelation of Almost Sure Invariance Principles for Certain Stochastic Adaptive Algorithms and for Partial Sums of Random Variables[1]

**Michael A. Kouritzin**[2]

*Received July 15, 1994; revised March 6, 1996*

Since the novel work of Berkes and Philipp[3] much effort has been focused on establishing almost sure invariance principles of the form

$$\left| \sum_{i=1}^{\lfloor t \rfloor} x_i - X_t \right| \ll t^{1/2-\gamma} \tag{1}$$

where $\{x_i, i=1,2,3,...\}$ is a sequence of random vectors and $\{X_t, t \geqslant 0\}$ is a Brownian motion. In this note, we show that if $\{A_k, k=1,2,3,...\}$ and $\{b_k, k=1,2,3,...\}$ are processes satisfying almost-sure bounds analogous to Eq. (1), (where $\{X_t, t \geqslant 0\}$ could be a more general Gauss–Markov process) then $\{h_k, k=1,2,3,...\}$, the solution of the stochastic approximation or adaptive filtering algorithm

$$h_{k+1} = h_k + \frac{1}{k}(b_k - A_k h_k) \qquad \text{for} \quad k=1,2,3,... \tag{2}$$

also satisfies an almost sure invariance principle of the same type.

**KEY WORDS:** Almost sure invariance principles; stochastic approximation; recursive algorithms; dependent random variables.

[1] This work was first presented in July 1993 at the Workshop über Mathematische Statistik sponsored by the Institut für Mathematische Stochastik, Universität Freiburg, Germany while the author was on an NSERC funded long-term visit there. The author is grateful to both NSERC and the aforementioned Institut for their support.

[2] Institute for Mathematics and Its Applications, University of Minnesota, Minneapolis, Minnesota 55455–0436.



860/9/4-1

# 1. INTRODUCTION

In 1951, Robbins and Monro[19] introduced their celebrated stochastic approximation procedure and suggested "the subject of stochastic approximation is likely to be useful and is worthy of future study." Evidently, they were right. Stochastic approximation procedures have enjoyed tremendous popularity in applications and analysis alike. Still, a completely satisfactory set of convergence results for the heavily-used, Robbins–Monro–based "linear adaptive-filtering algorithm":

$$h_{k+1} = h_k + \frac{1}{k}(b_k - A_k h_k) \qquad \text{for} \quad k = 1, 2, 3,... \tag{1.1}$$

(where $\{A_k,\ k = 1, 2, 3,...\}$ is a symmetric, positive semi-definite $\Re^{d \times d}$-valued process and $\{b_k, k = 1, 2, 3,...\}$ is a $\Re^d$-valued process) remains aloof. Many authors[8, 10, 12, 21] (also see their references) have proven results that imply central limit theorems, functional central limit theorems or laws of the iterated logarithm for Eq. (1.1) assuming $A_k$ (is constant or) converges almost surely and $\{b_k,\ k = 1, 2, 3,...\}$ is a martingale difference sequence. Unfortunately, neither the converging $A_k$ nor the martingale-difference assumptions hold in our applications. More recently, Ruppert,[20] Mark,[15] and Berger,[2] obtained almost sure invariance principles for Eq. (1.1) under more application-suitable assumptions on $\{b_k,\ k = 1, 2, 3,...\}$. However, $A_k(\omega)$ was still assumed (constant or) convergent and this previously appeared difficult to circumvent. In fact, in demonstrating that one can continue from Berger's work to attain an almost sure invariance principle for Eq. (1.1) when $\{A_k,\ k = 1, 2, 3,...\}$ and $\{b_k,\ k = 1, 2, 3,...\}$ are only mixing processes, Heunis[11] was forced to assume stringent-moment, strict-stationarity and certain strong or $\psi$-mixing conditions on $\{A_k, k = 1, 2, 3,...\}$ and $\{b_k,\ k = 1, 2, 3,...\}$. Weaker stationarity and mixing conditions are shown in Example 1 of our note to be somewhat unsuitable for many applications.

Having now intimated that an almost sure invariance principle for Eq. (1.1) under very general stationarity and dependence conditions would be of practical interest, we pose and partially solve a still more general problem. In this note, we assume that both $\{A_k,\ k = 1, 2, 3,...\}$ and $\{b_k, k = 1, 2, 3,...\}$ satisfy almost sure invariance principles to Gaussian processes $\{X_t^A, t \geq 0\}$ and $\{X_t^b, t \geq 0\}$ and ask what conditions on $\{(X^A \times X^b)_t, t \geq 0\}$ are sufficient to guarantee an almost sure invariance principle and a functional law of the iterated logarithm for $\{h_k,\ k = 1, 2, 3,...\}$. Wishing to keep our development manifest, we give our answer in terms of the solutions of simple integral equations. Because $\{A_k,\ k = 1, 2, 3,...\}$ and

$\{b_k,\ k=1,2,3,...\}$ are, in practice, often derived from the output of a possibly time-varying, nonexponentially-stable filter or from current parameter estimates, it is conceivable that the enlargement allowing $\{A_k,\ k=1,2,3,...\}$ and $\{b_k,\ k=1,2,3,...\}$ to satisfy invariance principles to certain Gauss–Markov processes other than Brownian motion, might assume more than just mathematical interest.

Past researchers, dating back to Walk,[21] have established their functional central limit theorems and invariance principles for Eq. (1.1) using

$$Y_t = W_t + (I-A)\int_0^1 \exp((A-2I^d)\log(\tau))\, W(\tau t)\, dt \qquad \text{for all} \quad t \geqslant 0 \qquad (1.2)$$

where $A-\frac{1}{2}I^d$ is a positive-definite matrix and $\{W_t,\ t\geqslant 0\}$ is a Brownian motion, as their approximating Gaussian process. We find it convenient to utilize an approximating Gaussian process with a slightly different form (evidenced by comparing Eq. (1.2) with the process in Remark 4 below). This new form facilitates easy association with the unique continuous solution of a simple stochastic integral equation as well as various simplifications throughout our proofs.

This note is organized as follows: Section 2 contains our notation, our conditions, and an expository application of the Algorithm in Eq. (1.1); our main result is stated at the beginning of Section 3; the proof of this result is housed in Sections 3–5; and some further examples are presented in Section 6.

## 2. NOTATION, CONDITIONS, AND EXAMPLE

Suppose $\{A_k^1,\ k=1,2,3,...\}$ is a symmetric, positive semi-definite, $\Re^{d\times d}$-valued stochastic process and $\{b_k^1,\ k=1,2,3,...\}$ is a $\Re^d$-valued stochastic process. Then, we consider the problem of establishing an almost sure invariance principle for $\{h_k^1,\ k=1,2,3,...\}$ generated by

$$h_{k+1}^1 = h_k^1 + \frac{1}{k}(b_k^1 - A_k^1 h_k^1) \qquad \text{for all} \quad k=1,2,3,... \qquad (2.1)$$

where $h_1^1$ is some nonrandom constant. Accordingly, assuming certain conditions on $\{A_k^1,\ k=1,2,3,...\}$ and $\{b_k^1,\ k=1,2,3,...\}$, we define a real constant $\delta>0$, a constant vector $h_*$, and a Gaussian process $\{Y_t,\ t\geqslant 0\}$ on a second probability space $(\Omega,\mathscr{F},P)$ and prove that there exists a process $\{h_k,\ k=1,2,3,...\}$ on $(\Omega,\mathscr{F},P)$ such that (i) $\{h_k,\ k=1,2,...\} \stackrel{\mathscr{D}}{=} \{h_k^1,\ k=1,2,...\}$ and (ii) $|\lfloor t\rfloor(h_{\lfloor t+1\rfloor}-h_*)-Y_t| \ll^t t^{1/2-\delta}$ for all $t>0$ a.s.

Striving for generality, we assume only that $\{A_k^1, k=1,2,3,...\}$ and $\{b_k^1, k=1,2,3,...\}$ themselves satisfy almost sure invariance principles.

To be more specific, we will conform to the following notation list:

$|x|$ is the Euclidean distance of a $\mathfrak{R}^k$-vector $x$ (any $k$) and $|||A||| = \sup_{|x|=1} |Ax|$ for a $\mathfrak{R}^{k\times k}$-matrix $A$.

$\{z_k, k=1,2,3,...\} \overset{\mathscr{D}}{=} \{\tilde{z}_k, k=1,2,3,...\}$ means equal in distribution whereas $\{Y_t, t\geqslant 0\} = \{Z_t, t\geqslant 0\}$ means $\{Y_t, t\geqslant 0\}$ and $\{Z_t, t\geqslant 0\}$ are indistinguishable.

$\lfloor t\rfloor \triangleq \max\{i\in\mathbb{N}: i\leqslant t\}$, $\lceil t\rceil \triangleq \min\{i\in\mathbb{N}: i\geqslant t\}$, and $a\vee p$ $(a\wedge p) =$ maximum (minimum) of $a, p$.

$a_{i,k} \ll^{i,k} b_{i,k}$ implies there is a $c>0$ not depending on $i$ or $k$ such that $|a_{i,k}| \leqslant c\,|b_{i,k}|$ for all $i, k$.

$s^A = \exp(A\cdot\log(s))$ when $A$ is a square matrix and $s>0$.

$I^d$, $0^d = d\times d$ identity, zero matrices and $\prod_{l=p}^{q} B_l$ (matrix product) $= B_q B_{q-1}\cdots B_p$ or $I^d$ if $p>q$.

$\mathbf{var}(Y) = E\{(Y-EY)(Y-EY)^T\}$ for any $\mathfrak{R}^d$-valued random vector.

$A^{(n)} = n$th column of $A$ or $n$th element of $A$ if $A$ is a vector, $A^{(n,m)} = (n,m)$th component of $A$, and $\sigma^2_{(n,n)}(Y) = (\mathrm{var}(Y))^{(n,n)}$.

$\delta_{n,l}$ is Kronecker delta and $1(a,b,c)$ equals 1 if $a\leqslant b\leqslant c$ and otherwise it is equal to 0.

Then, we hypothesize existence of a certain $\mathfrak{R}^{d\times d}\times\mathfrak{R}^d$-valued Gaussian process $\{(X^A\times X^b)_t, t\geqslant 0\}$ and processes $\{A_k, k=1,2,3,...\}$, $\{b_k, k=1,2,3,...\}$ on $(\Omega, \mathscr{F}, P)$ such that: (a) $\{A_k, k=1,2,3,...\} \overset{\mathscr{D}}{=} \{A_k^1, k=1,2,3,...\}$, (b) $\{b_k, k=1,2,3,...\} \overset{\mathscr{D}}{=} \{b_k^1, k=1,2,3,...\}$, and (c) for some $0<\gamma\leqslant\frac{1}{2}$ and (unknown) $A_*$, $b_*$ we have that

$$\left|\left|\left| \sum_{k\leqslant\lfloor t\rfloor} (A_k - A_*) - X_t^A \right|\right|\right| \ll^t t^{1/2-\gamma}$$

and

$$\left| \sum_{k\leqslant\lfloor t\rfloor} (b_k - b_*) - X_t^b \right| \ll^t t^{1/2-\gamma} \qquad \text{for all} \quad t>0 \qquad \text{a.s.}$$

Next, letting $h_* \triangleq A_*^{-1} b_*$, $z_k \triangleq b_k - A_k h_* = b_k - b_* - (A_k - A_*)\,h_*$, and

$$v_{k+1} = v_k + \frac{1}{k}(z_k - A_k v_k) \qquad \text{for} \quad k=1,2,3,... \tag{2.2}$$

subject to $v_1 \triangleq h_1^1 - h_*$, we see immediately from Eq. (2.1) and Eq. (2.2) that our initial problem has been reduced to defining $\{Y_t, t \geqslant 0\}$ on $(\Omega, \mathscr{F}, P)$ and showing that for some $\delta > 0$

$$|\lfloor t \rfloor v_{\lfloor t+1 \rfloor} - Y_t| \ll^t t^{1/2-\delta} \qquad \text{for all} \quad t > 0 \qquad \text{a.s. } [P] \qquad (2.3)$$

Our approach does not require $(X^A \times X^b)_t, t \geqslant 0\}$ to be a Brownian motion but rather only presumes that it satisfies three properties (which will be motivated in Section 6). Aiming to keep our development brief and readily discernible, we describe these properties in terms of a class of functions $F: [0, \infty) \to \mathfrak{R}^{(d^2+d) \times (d^2+d)}$ such that $F(0) = 0$ and each component of $F$ is continuous and has finite variation on compact intervals. Then, the total variation of each component of $F$ must also be a uniformly continuous function on compact intervals so by dividing an arbitrary interval $[0, \tau]$, $\tau > 0$ into small enough sub-intervals it follows from a fixed point theorem that

$$G_t = I^{d^2+d} + \int_0^t dF_s \cdot G_s \qquad (2.4)$$

has a unique continuous solution, $\mathscr{E}^R(F)$ (the right exponential of $F$), on $[0, \tau]$ and hence on $[0, \infty)$. Next, being bounded on compact intervals and satisfying Eq. (2.4), $\mathscr{E}^R(F)$ must have finite variation on compact intervals. Finally, it then follows by analogous results (for Riemann–Stieltjes integration) to Protter[18] [Thm. 48, p. 264, and Thm. 19, p. 55] that $\mathscr{E}^R(F)_t$ is nonsingular for each $t \geqslant 0$ and its inverse is also continuous. Letting

$$\mathscr{E}_F(t, s) \triangleq \mathscr{E}^R(F)_t \, (\mathscr{E}^R(F)_s)^{-1} \qquad \text{for all} \quad 0 \leqslant s \leqslant t \qquad (2.5)$$

we define $\mathscr{M}$ to be the class of such continuous, finite-variation $F$ such that:

(i) $$\sup_{0 \leqslant s \leqslant t \leqslant s+1} |||\mathscr{E}_F^T(t, s)||| < \infty$$

(ii) $$\sum_{p=1}^{q} \sup_{u \geqslant 0} |||\mathscr{E}_F^T(\lceil u \rceil + p, u) - \mathscr{E}_F^T(\lceil u \rceil + p - 1, u)||| \ll^q 1 + \log(q)$$

(iii) $$\int_0^q \sup_{u \geqslant 0} |||\mathscr{E}_F^T(\lceil u \rceil + v, u) - \mathscr{E}_F^T(\lceil u \rceil + \lfloor v \rfloor, u)||| \, dv \ll^q 1 + \log(q)$$

for all integers $q \geqslant 1$ and impose our two basic hypotheses:

**(C1).** Suppose $\{X_t,\ t\geqslant 0\}$ is the pathwise unique continuous solution of

$$X_t = BW_t + \int_0^t dF_s \cdot X_s$$

where $F$ is a ($\Re^{(d^2+d)\times(d^2+d)}$-valued) class $\mathscr{M}$ function, $\{W_t,\ t\geqslant 0\}$ is a $\Re^e$-valued Brownian motion on $(\Omega, \mathscr{F}, P)$ with $W_0 \equiv 0$, and $B$ is a $\Re^{(d^2+d)\times e}$-matrix. Then, (i) letting $\{X_t^z,\ t\geqslant 0\}$ denote the vector process consisting of the first $d$ elements of $\{X_t,\ t\geqslant 0\}$, we assume that

$$\left|\sum_{r\leqslant \lfloor t\rfloor} z_r - X_t^z\right| \ll^t t^{1/2-\gamma} \qquad \text{for all} \quad t>0 \qquad \text{a.s. } [P]$$

and (ii) letting $\{X_t^A,\ t\geqslant 0\}$ denote the $\Re^{d\times d}$-valued process whose columns consist of the remaining elements of $\{X_t,\ t\geqslant 0\}$ taken in order and $d$ at a time, we assume that

$$\left|\left|\left|\sum_{r\leqslant \lfloor t\rfloor} \bar{A}_r - X_t^A\right|\right|\right| \ll^t t^{1/2-\gamma} \qquad \text{for all} \quad t>0 \qquad \text{a.s. } [P]$$

where $\bar{A}_r \triangleq A_r - A_*$ for $r = 1, 2, 3,\ldots$

**(C2).** The smallest eigenvalue of $A_*$, $\lambda_{\min}$, in Condition (C1) satisfies $\lambda_{\min} > \frac{1}{2}$.

**Remark 1.** All the stochastic integral equations in this note are of the form:

$$X_t^{(i)} = J_t^{(i)} + \sum_{j=1}^{k} \int_0^t X_s^{(j)}\, dF_s^{(i,j)} \qquad i = 1, 2,\ldots, k \quad \text{or} \quad X_t = J_t + \int_0^t dF_s \cdot X_s$$

$$\text{for all} \quad t\geqslant 0$$

where the components of $F$ are continuous functions with finite variation on compact intervals and $\{J_t,\ t\geqslant 0\}$ is a given continuous-path semimartingale. The argument previously used to define the right exponential of $F$ can be adapted to establish that for each $\omega\in\Omega$ this integral equation has a unique continuous solution $X(\omega)$ on $[0, \infty)$. Furthermore, the "variation of constants" method (see e.g., Protter[18], [p. 267]), and Thms. 13 and 19 of Protter[18] [p. 53, 55] can be used to establish

$$\{X_t,\ t\geqslant 0\} = \left\{\int_0^t \mathscr{E}_F(t, s)\, dJ_s,\ t\geqslant 0\right\}$$

Hence, $\{X_t,\ t \geq 0\}$ is a continuous semimartingale indistinguishable from $\{\int_0^t \mathscr{E}_F(t, s)\, dJ_s,\ t \geq 0\}$ and specifying that $F$ is class $\mathscr{M}$ is our method of limiting the behavior of $\{X_t,\ t \geq 0\}$. Indeed, we have stipulated that a class $\mathscr{M}$ function should satisfy (i) and (ii) above largely to obtain a convenient bound on the variance of $X_r - X_{r-1}$ (for $r = 1, 2,...$) when $\{J_t,\ t \geq 0\}$ is a Brownian motion.

**Remark 2.** It is easy to see that (C1) is equivalent to almost sure invariance principles for $\{b_r - b_*,\ r = 1, 2, 3,...\}$ and $\{\bar{A}_r,\ r = 1, 2, 3,...\}$. Moreover, since $h_*$ is never known when one uses Eq. (2.1) and (C1) (i) can be restated as

$$\left| \sum_{r \leq \lfloor t \rfloor} ((b_r - b_*) - \bar{A}_r h_*) - X_t^{\pm} \right| \ll^t t^{1/2-\gamma} \qquad \text{for all} \quad t > 0 \qquad \text{a.s. } [P]$$

there is, at least from a pragmatic point of view, no loss of generality in imposing (C1) (ii) once (C1) (i) has been assumed. Furthermore, as illustrated in Example 1, $\{b_r,\ r = 1, 2, 3,...\}$ and $\{A_r,\ r = 1, 2, 3,...\}$ invariably contain the same type of data so this dual invariance principle assumption is a natural one.

**Example 1.** Consider second-order $\{y_k,\ k = 1, 2, 3,...\}$ and $\{u_k,\ k = 1, 2, 3,...\}$ satisfying

$$y_{k+1}(\omega) = h^{(1)} y_k(\omega) + h^{(2)} u_k(\omega) \qquad \text{for} \quad k = 1, 2, 3,... \tag{2.6}$$

where $h \in \Re^2$ is a vector of unknown parameters with $0 < |h^{(1)}| < 1$. Now, suppose that we only have access to corrupted data defined by $\psi_k(\omega) = y_k(\omega) + \rho_k(\omega)$ and $e_k(\omega) = u_k(\omega) + \zeta_k(\omega)$, where $\{\rho_k,\ k = 1, 2, 3,...\}$ and $\{\zeta_k,\ k = 1, 2, 3,...\}$ are zero-mean, second-order, i.i.d. random variables on $(\Omega^1, \mathscr{F}^1, P^1)$ independent of each other, $y_1$ and $\{u_k,\ k = 1, 2, 3,...\}$. Then, to estimate $h$ recursively one often uses Algorithm in Eq. (2.1) with

$$b_k^1 = \begin{bmatrix} \psi_{k+1} \psi_k \\ \psi_{k+1} e_k \end{bmatrix}, \qquad A_k^1 = \begin{bmatrix} \psi_k^2 & \psi_k e_k \\ \psi_k e_k & e_k^2 \end{bmatrix} \qquad \text{for} \quad k = 1, 2, 3,... \tag{2.7}$$

$A_k^1(\omega)$ is symmetric and positive semi-definite for all $k$ and $\omega$ but not normally convergent as $k \to \infty$. Next, to show that weak-stationarity conditions for $\{b_k^1 - Eb_k^1,\ k = 1, 2, 3,...\}$ and $\{A_k^1 - EA_k^1,\ k = 1, 2, 3,...\}$ often fail (without introducing horrendous calculations), we assume that $y_1 = 0$

and $\{u_k, k=1,2,3,...\}$ is a zero-mean, i.i.d. sequence with $Eu_k^2=1$ and find by Eq. (2.6) that for all $k=1,2,3,...$

$$E\{(\psi_k e_k - E\{\psi_k e_k\})^2\}$$
$$= E\{\psi_k^2\}\, E\{e_k^2\} = \left[\frac{(h^{(2)})^2}{1-(h^{(1)})^2}(1-(h^{(1)})^{2k-2}) + E\rho_1^2\right][1+E\zeta_1^2] \qquad (2.8)$$

which depends on $k$. On a different note, a glance at Eq. (2.7) leads one to believe that $\{A_k^1, k=1,2,3,...\}$ would normally satisfy conditions akin to those for $\{b_k^1, k=1,2,3,...\}$ which is, in fact, our motivation for Condition (C1) (ii). However, even standard mixing conditions may be less-than-ideal assumptions when these processes contain discrete data. To illustrate this suppose $\{u_k, k=1,2,3,...\}$ is a (not-necessarily-independent) identically-distributed sequence such that $P^1(u_k=2)=P^1(u_k=4)=\frac{1}{2}$; $y_1$ is such that $P^1(y_1=\frac{1}{3})=P^1(y_1=\frac{2}{3})=\frac{1}{2}$ and $h^{(1)}=h^{(2)}=\frac{1}{6}$; $P^1(|\zeta_k|\geqslant 2)=0$ for all $k$; and $P^1(\rho_k=\frac{1}{36})=P^1(\rho_k=-\frac{1}{36})=\frac{1}{2}$ for all $k$. Then, it follows easily from Eq. (2.6) that

$$\psi_k = \frac{1}{6}u_{k-1} + \frac{1}{36}u_{k-2} + \frac{1}{216}u_{k-3} + \cdots + \frac{1}{6^{k-1}}u_1 + \frac{1}{6^{k-1}}y_1 + \rho_k \qquad (2.9)$$

for each $k$, $\omega$ and letting $\mathscr{A}_k$ denote the $\sigma$-algebra of sets of the form $(A\cap\tilde{\Omega})\cup N$ for any $A\in\sigma(\psi_k)$, $P^1$-unit set $\tilde{\Omega}$, and $P^1$-null set $N$; we find by Eq. (2.9) that

$$\mathscr{A}_k \supset \sigma(y_1, u_1, u_2,..., u_{k-3}, u_{k-2}+\rho_k, u_{k-1}) \supset \sigma(y_1) \quad \text{for all} \quad k=4,5,6,... \qquad (2.10)$$

Finally, noting that

$$\psi_k = |\psi_k|\,\mathrm{sgn}(\psi_k) = (\psi_k^2)^{1/2}\,\mathrm{sgn}(\psi_k e_k) \qquad \text{for all} \quad k=1,2,3,... \quad \text{a.s.} \qquad (2.11)$$

we can use Eqs. (2.10) and (2.11), and arguments similar to Bradley[4] [p. 180] to conclude $\{(\psi_k^2)^{1/2}\,\mathrm{sgn}(\psi_k e_k), k=1,2,3,...\}$, whence $\{A_k^1, k=1,2,3,...\}$, fails to be strong mixing (with any mixing rate). Our conclusions from this example appear to remain important when considering more involved problems in adaptive equalization, adaptive array processing, linear classification and ARMA modeling (see Widrow and Stearns[23] [Chap. 6], Benveniste *et al.*[1] [Chap. 1], and the introduction and references of Farden[9] for further examples and discussion). Therefore, we have formed our hypotheses with these conclusions in mind and we will begin to illustrate the generality of these conditions in Section 6.

We would like to thank a referee for pointing out that Walk,[22] in an apparently little known result published outside of the standard probability journals, independently discovered our idea of transferring an almost sure invariance principle from the data processes to the filter output $\{h_k, k=1,2,...\}$. However, there are some significant differences between our work and Walk's. Most notably, Walk does not consider invariance to processes other than Brownian motion, he only attempts a loglog invariance principle *vis-à-vis* our tighter form of Gaussian approximation, and his method is entirely different. Still, it is the author's opinion that Walk's work is a very significant advancement.

In the next section we state and prove our desired almost sure invariance principle for $\{v_k, k=1,2,3,...\}$ defined in Eq. (2.2), assuming we already have one for $\{u_k, k=1,2,3,...\}$ defined by

$$u_1 = h_1^! - h_* \qquad \text{and} \qquad u_{k+1} = u_k + \frac{1}{k}(z_k - A_* u_k)$$
$$\text{for all} \quad k = 1, 2, 3,... \tag{2.12}$$

This second almost sure invariance principle is the subject of Section 4.

## 3. AN ALMOST SURE INVARIANCE PRINCIPLE FOR $\{v_k, k=1,2,3,...\}$

**Theorem 1.** Under Conditions (C1) and (C2) of Section 2 one has that:

$$|\lfloor t \rfloor v_{\lfloor t+1 \rfloor} - Y_t| \ll^t t^{1/2-\delta} \qquad \text{for all} \quad t \geqslant 0 \qquad \text{a.s. } [P],$$

for some $\delta > 0$, where $Y_t \triangleq \int_0^t ((s+1)/(t+1))^{A_* - I} \, dX_s^z$ for all $t \geqslant 0$.

**Remark 3.** In our proof, we will demonstrate that $\delta$ is arbitrary, provided $\delta \leqslant \gamma/8$ and $\delta < \lambda_{\min} - \frac{1}{2}$, $\lambda_{\min}$ and $\gamma$ being the constants of Conditions (C1) and (C2).

**Remark 4.** Applying Itô's formula to $\{(t+1)^{A_* - I} Z_t, t \geqslant 0\}$, where $\{Z_t, t \geqslant 0\}$ solves

$$Z_t = \int_0^t \frac{I - A_*}{s+1} Z_s \, ds + X_s^z \qquad \text{for all} \quad t \geqslant 0$$

one finds that $\{Z_t, t \geqslant 0\} = \{Y_t, t \geqslant 0\}$ so $\{Y_t, t \geqslant 0\}$ is the almost-surely unique continuous solution to this linear stochastic integral equation. This fact will be used in Example 3. Now, if we apply integration by parts to $\{\int_0^t ((s+1)/(t+1))^{A_*-I}\, dX_s^z,\ t\geqslant 0\}$, we find

$$\{Y_t,\ t\geqslant 0\}=\left\{X_t^z-\int_0^t \frac{A_*-I}{t+1}\left(\frac{s+1}{t+1}\right)^{A_*-2I} X_s^z\, ds,\ t\geqslant 0\right\}$$

which will be utilized in Section 5.

**Remark 5.** To utilize the stability of Eq. (2.2) imposed by Conditions (C1) (ii) and (C2); we let (a) $n_i \triangleq \lfloor (i\gamma/2)^{2/\gamma} \rfloor$ and (b) $I_i \triangleq \{n_i+1, n_i+2,..., n_{i+1}\}$ for each $i=0, 1, 2,...$; and (c) $s_k$ be the unique index such that $k-1\in I_{s_k}$ for $k=2, 3, 4,...$ It then follows easily that (d) $\sum_{j\in I_i} j^{-1}\leqslant (n_i+1)^{-\gamma/2}\sum_{j\in I_i} j^{\gamma/2-1}\leqslant (n_i+1)^{-\gamma/2}+(n_i+1)^{-1}$ and (e) $(n_{i+1}+1)\leqslant (2^{2/\gamma}+1)(n_i+1)$ for $i=0, 1, 2,...$

**Remark 6.** It follows (see Ref. 13) from this theorem that $V^N(\tau)\triangleq \lfloor \tau N\rfloor v_{\lfloor \tau N+1\rfloor}$ for all $\tau\in[0, 1]$ and $N=3, 4, 5,...$ satisfies a functional law of the iterated logarithm if $\Phi^N(\tau)\triangleq X_{\tau N}^z$ does.

### 3.1. Proof of Theorem 1

We assume ($e=d^2+d$ and) $B=I^{d^2+d}$ throughout the sequel.

*Proof.* By Proposition 1 (i) and Lemma 4 (b) below, it only remains to show that there exists a $\kappa>0$ such that for a.a. $\omega$

$$|w_{\lfloor t+1\rfloor}|\ll^t t^{-1/2-\kappa} \qquad \text{for all} \quad t\geqslant 1 \tag{3.1}$$

where $w_k\triangleq v_k-u_k$ for $k=1, 2, 3,...$ However, by Condition (C1), and Remark 1 it follows that

$$X_t=\int_0^t \mathscr{E}(t, s)\, dW_s \qquad \text{for all} \quad t\geqslant 0 \qquad \text{a.s.} \tag{3.2}$$

where $\mathscr{E}(t, s)\triangleq \mathscr{E}_F(t, s)$ is defined in Eq. (2.5). Hence, clearly $\{X_t,\ t\geqslant 0\}$ is a zero-mean Gaussian process, and for each $r=1, 2, 3,...$ and $n=1, 2,..., d^2+d$ one has that

$$\sigma^2_{(n,n)}(X_r-X_{r-1})=\int_{r-1}^r |\mathscr{E}^T(r, s)^{(n)}|^2\, ds+\int_0^{r-1} |\mathscr{E}^T(r, s)^{(n)} - \mathscr{E}^T(r-1, s)^{(n)}|^2\, ds \tag{3.3}$$

by independence and the usual Hilbert space isometry. Therefore, by Lemma 2 (with $p=q=r$ and $G_p=(I^{d^2+d})^{(n)})$ and Cauchy–Schwarz we have a $c_1>0$ such that

$$|||\mathrm{var}(X_r-X_{r-1})||| \leqslant c_1(1+\log(r))^2 \qquad \text{for all} \quad r=1,2,3,... \tag{3.4}$$

and so for a.a. $\omega$ $[P]$ we have (by Borel–Cantelli) that

$$|||X_r^A-X_{r-1}^A|||+|X_r^z-X_{r-1}^z| \ll^r 1+\log^2(r) \qquad \text{for all} \quad r=1,2,3,... \tag{3.5}$$

Furthermore, fixing an integer $k\geqslant 2$ and defining

$$\tilde{A}_j \triangleq \bar{A}_j-X_j^A+X_{j-1}^A, \qquad \tilde{z}_j \triangleq z_j-X_j^z+X_{j-1}^z \qquad \text{for all} \quad j=1,2,3,... \tag{3.6}$$

we have by Eq. (3.5) and Condition (C1) (ii) that the largest eigenvalue of $A_k$ satisfies

$$\frac{1}{k}\lambda_{\max}(A_k) \leqslant \frac{1}{k}\{|||X_k^A-X_{k-1}^A|||+|||\tilde{A}_k|||+|||A_*|||\} \ll^k k^{-1/2-\gamma} \qquad \text{a.s.} \tag{3.7}$$

Hence, (recalling $w_1=0$ and $w_{j+1}=v_{j+1}-u_{j+1}=(I-(1/j)\,A_j)\,w_j-(1/j)\,\bar{A}_j u_j$ for $j=1,2,3,...$) we have by Eqs. (2.2), (2.12), (3.6), Proposition 1 (ii), Eqs. (3.7) and (3.5) that

$$\begin{aligned} |w_k| &\ll^k \sum_{l=1}^{k-1}\prod_{j=l+1}^{k-1}\left|\left|\left|I-\frac{1}{j}A_j\right|\right|\right|\cdot l^{-3/2}(1+\log(l))^2\{|||X_l^A-X_{l-1}^A|||+|||\tilde{A}_l|||\} \\ &\ll^k \prod_{j:\,\lambda_{\max}(A_j)\geqslant 2j}\left(\frac{\lambda_{\max}(A_j)}{j}-1\right)\cdot\sum_{l=1}^{k-1} l^{-1-\gamma/2} \ll^k 1 \qquad \text{a.s.} \end{aligned} \tag{3.8}$$

The result will follow by induction if given any $-\frac{1}{2}<\eta\leqslant 0$ we can show that

$$|w_k| \ll^k k^\eta \qquad k=1,2,3,... \qquad \text{a.s.} \tag{3.9}$$

can be refined to $|w_k| \ll^k k^\rho$ for all $k=2,3,4,...$, where $\rho<0$ is arbitrary provided $\rho\geqslant\max\{\eta-\frac{1}{2}+\gamma/2,\ -\frac{1}{2}-\gamma/8\}$ and $\rho>-\lambda_{\min}$. First, we fix an integer $k\geqslant 2$, and define

$$F_{j,k} \triangleq \prod_{r=j+1}^{k-1}\left(I-\frac{1}{r}A_*\right)\frac{1}{j} \qquad \text{for} \quad j=1,2,...,k-1 \tag{3.10}$$

Then, symmetry, positive semi-definiteness, and Eq. (3.6) imply

$$\sum_{n \in I_i} n^{\eta-1} |||A_n||| \ll^{i,k} \left|\left|\left| \sum_{n \in I_i} n^{\eta-1} A_n \right|\right|\right|$$

$$\ll^{i,k} \sum_{n \in I_i} n^{\eta-1} (|||A_*||| + |||X_n^A - X_{n-1}^A|||)$$

$$+ \left|\left|\left| \sum_{n \in I_i} n^{\eta-1} \tilde{A}_n \right|\right|\right| \tag{3.11}$$

$$\sum_{j \in I_i} |||\bar{A}_j||| \ll^{i,k} \sum_{j \in I_i} |||A_*||| + \left|\left|\left| \sum_{j \in I_i} A_j \right|\right|\right|$$

$$\ll^{i,k} \sum_{j \in I_i} (|||A_*||| + |||X_j^A - X_{j-1}^A|||) + \left|\left|\left| \sum_{j \in I_i} \tilde{A}_j \right|\right|\right| \tag{3.12}$$

for $i = 0, 1, ..., s_k$ and a.a. $\omega$. Hence, it follows from Eq. (2.2), (2.12), (3.10), (3.9), and Proposition 1 (ii) that

$$|w_k| \leqslant \sum_{i=0}^{s_k} \left\{ \left|\left|\left| \sum_{\substack{j \in I_i \\ j<k}} F_{j,k} \bar{A}_j \right|\right|\right| \cdot |v_{n_i+1}| + \left| \sum_{\substack{j,n \in I_i \\ n<j<k}} F_{j,k} \bar{A}_j \frac{1}{n} z_n \right| \right.$$

$$\left. + \left| \sum_{\substack{j,n \in I_i \\ n<j<k}} F_{j,k} \bar{A}_j \frac{1}{n} A_n v_n \right| \right\}$$

$$\ll^k \sum_{i=0}^{s_k} \left\{ \left|\left|\left| \sum_{\substack{j \in I_i \\ j<k}} F_{j,k} \bar{A}_j \right|\right|\right| \cdot (n_i+1)^\eta + \sum_{\substack{n \in I_i \\ n<k}} \left[ \left|\left|\left| \sum_{\substack{j \in I_i \\ n<j<k}} F_{j,k} \bar{A}_j \right|\right|\right| \right. \right.$$

$$\left. \times \frac{1}{n} (|X_n^z - X_{n-1}^z| + |||A_n||| \, n^\eta) \right]$$

$$\left. + \sum_{\substack{j \in I_i \\ j<k}} \left[ |||F_{j,k}||| \cdot |||\bar{A}_j||| \cdot \left| \sum_{\substack{n \in I_i \\ n<j}} \frac{1}{n} \tilde{z}_n \right| \right] \right\} \tag{3.13}$$

so one has by Eqs. (3.5), (3.11), (3.12), Lemma 1 (a–e) and Remark 5 (a–e) that

$$|w_k| \ll^k \sum_{i=0}^{s_k} \left\{ \max_{\substack{p,q \in I_i \\ p \leq q < k}} \left\| \sum_{j=p}^{q} F_{j,k} \bar{A}_j \right\| \cdot \left[ (n_i+1)^\eta \right.\right.$$

$$+ \sum_{\substack{n \in I_i \\ n<k}} \frac{1+\log^2(n)}{n} (1+n^\eta) + \left. \left\| \sum_{n \in I_i} n^{\eta-1} \tilde{A}_n \right\| \right]$$

$$\left. + \max_{\substack{m \in I_i \\ m<k}} \left[ \|F_{m,k}\| \cdot \left| \sum_{\substack{n \in I_i \\ n<m}} \frac{1}{n} \tilde{z}_n \right| \right] \cdot \left[ \sum_{j \in I_i} (1+\log^2(j)) + \left\| \sum_{j \in I_i} \tilde{A}_j \right\| \right] \right\}$$

$$\ll^k \sum_{i=0}^{s_k} \{(1+\log^2(n_{s_k}+1))(n_{s_k}+1)^{-\lambda_{\min}} (n_i+1)^{\lambda_{\min}-1/2-\gamma/4}$$

$$\times [(n_i+1)^\eta + (1+\log^2(n_i+1))(n_i+1)^{-\gamma/2}]$$

$$+ (n_{s_k}+1)^{-\lambda_{\min}} (n_i+1)^{\lambda_{\min}-3/2-\gamma} [(1+\log^2(n_i+1))$$

$$\times (n_{i+1}-n_i) + (n_i+1)^{1/2-\gamma}]\}$$

$$\ll^k (n_{s_k}+1)^{-\lambda_{\min}} (1+\log^2(n_{s_k}+1)) \sum_{i=0}^{s_k} \{(n_i+1)^{\lambda_{\min}-1/2-\gamma/4}$$

$$\times [(n_i+1)^\eta + (n_i+1)^{-7\gamma/16}] + (n_i+1)^{\lambda_{\min}-1/2-3\gamma/4}\}$$

$$\ll^k k^{-\lambda_{\min}} \log^2(k) \cdot \left[ 1 + \sum_{i=1}^{s_k} \{ i^{2/\gamma(\lambda_{\min}-1/2-\gamma/4+\eta)} + i^{2/\gamma(\lambda_{\min}-1/2-11/16\gamma)} \} \right]$$

$$\ll^k k^{-\lambda_{\min}} \log^2(k) \cdot [1 + (n_{s_k}+1)^{\lambda_{\min}-1/2+\gamma/4+\eta} + (n_{s_k}+1)^{\lambda_{\min}-1/2-3\gamma/16}]$$

$$\ll^k k^{-\lambda_{\min}} \log^2(k) + k^{-1/2+\gamma/2+\eta} + k^{-1/2-\gamma/8}$$

$$\text{for all} \quad k=2,3,4,\ldots \quad \text{and} \quad \text{a.a. } \omega \tag{3.14}$$

provided $2/\gamma(\lambda_{\min}-1/2-\gamma/4+\eta) \neq -1$ and $2/\gamma(\lambda_{\min}-1/2-\gamma 11/16) \neq -1$. These cases are handled with trivial modifications to this procedure and our proof is therefore complete. □

### 3.2. Technical Bounds Required for Theorem 1

**Lemma 1.** Suppose $\tilde{A}_j$ and $\tilde{z}_j$ are as defined in Eq. (3.6), $F_{m,k}$ is as in Eq. (3.10) and Conditions (C1) and (C2) of Section 2 hold. Then, it follows for a.a. $[P]$ $\omega$, all integers $k \geq 2$ and all $i \in \{0, 1,..., s_k\}$ that

(a) $\max_{\substack{p, q \in I_i \\ p \leqslant q}} \left|\left|\left| \sum_{j=p}^{q} \tilde{A}_j \right|\right|\right| \ll^{i,k} (n_i + 1)^{1/2 - \gamma}$

(b) $\max_{\substack{p, q \in I_i \\ p \leqslant q}} \left| \sum_{j=p}^{q} \frac{1}{j} \tilde{z}_j \right| \ll^{i,k} (n_i + 1)^{-1/2 - \gamma}$

(c) $\left|\left|\left| \sum_{j \in I_i} j^{\eta - 1} \tilde{A}_j \right|\right|\right| \ll^{i,k} (n_i + 1)^{\eta}$

(d) $\max_{\substack{m \in I_i \\ m < k}} |||F_{m,k}||| \ll^{i,k} (n_i + 1)^{\lambda_{\min} - 1} (n_{s_k} + 1)^{-\lambda_{\min}}$

(e) $\max_{\substack{p, q \in I_i \\ p \leqslant q < k}} \left|\left|\left| \sum_{j=p}^{q} F_{j,k} \tilde{A}_j \right|\right|\right| \ll^{i,k} (n_{s_k} + 1)^{-\lambda_{\min}}$

$\times (1 + \log^2(n_{s_k} + 1))(n_i + 1)^{\lambda_{\min} - 1/2 - \gamma/4}$

*Proof.* Insomuch as arguments for (c) and (e) can be modified for (a), (b) and (d), the latter three proofs are omitted. Suppose $i, k$ are such that $k \geqslant 2$, $0 \leqslant i \leqslant s_k$ and $I_i$ is not empty.

(c) We have by the substitution $j^{\eta - 1} = (n_i + 1)^{\eta - 1} + \sum_{r \in I_i, r < j} ((r + 1)^{\eta - 1} - r^{\eta - 1})$, an interchange of summation, the mean value theorem, Condition (C2) (i), and Remark 5 (d, e) that

$$\left|\left|\left| \sum_{j \in I_i} j^{\eta - 1} \tilde{A}_j \right|\right|\right| \ll^{i,k} \left[ (n_i + 1)^{\eta - 1} + \sum_{r \in I_i} r^{\eta - 2} \right] \left\{ \max_{\substack{p, q \in I_i \\ p \leqslant q}} \left|\left|\left| \sum_{j=p}^{q} \tilde{A}_j \right|\right|\right| \right\}$$

$$\ll^{i,k} (n_i + 1)^{\eta - 1/2 - \gamma} \qquad \text{a.s. } [P] \tag{3.15}$$

(e) Using (a) and partial summation of $F_{j,k}$, we have for any $p, q \in I_i$ such that $p \leqslant q < k$ that

$$\left|\left|\left| \sum_{j=p}^{q} F_{j,k} \tilde{A}_j \right|\right|\right| \ll^{i,k,p,q} (n_i + 1)^{1/2 - \gamma} \left\{ \max_{\substack{m \in I_i \\ m < k}} |||F_{m,k}||| \right.$$

$$\left. + \sum_{\substack{r \in I_i \\ r < k - 1}} |||F_{r+1,k} - F_{r,k}||| \right\} \qquad \text{a.s. } [P] \tag{3.16}$$

Moreover, letting $\lambda$ denote an arbitrarily chosen eigenvalue of $A_*$, noting that

$$\prod_{l=r+2}^{k-1} \left|1-\frac{\lambda}{l}\right| \leqslant \left[\prod_{l=2}^{\lfloor \lambda/2 \rfloor} \left(\frac{\lambda}{l}-1\right) \cdot \exp\left\{\sum_{l=2}^{\lfloor \lambda \rfloor - 1} \frac{\lambda}{l}\right\}\right] \times \exp\left\{-\lambda \int_{r+1}^{k-1} \frac{dt}{t}\right\} \quad \text{for all} \quad r=0, 1, 2,... \tag{3.17}$$

and recalling from Remark 5 that $k-1 \geqslant n_{s_k}+1$ and $n_{i+1} \ll^i n_i + 1$, we find that

$$\left|\prod_{l=r+2}^{k-1} \left(1-\frac{\lambda}{l}\right) \frac{1}{r+1} - \prod_{l=r+1}^{k-1} \left(1-\frac{\lambda}{l}\right) \frac{1}{r}\right| \ll^{r,i,k} (n_{s_k}+1)^{-\lambda_{\min}} (n_i+1)^{\lambda_{\min}-1} \frac{1}{r} \tag{3.18}$$

for $r = n_i + 1,..., n_{i+1} - 1$ so by Eqs. (3.6), (3.16), (d), (3.10), the fact that the eigenvectors of $A_*$ span $\mathfrak{R}^d$, the principle of uniform boundedness, Eq. (3.18) and Remark 5 (d) it follows for a.a. $\omega$ that

$$\max_{\substack{p,q \in I_i \\ p \leqslant q < k}} \left|\left|\left| \sum_{j=p}^{q} F_{j,k} \bar{A}_j \right|\right|\right| \ll^{i,k} (n_{s_k}+1)^{-\lambda_{\min}} (n_i+1)^{\lambda_{\min}-1/2-\gamma} + \max_{\substack{p,q \in I_i \\ p \leqslant q < k}} \left|\left|\left| \sum_{j=p}^{q} F_{j,k}(X_j^A - X_{j-1}^A) \right|\right|\right| \tag{3.19}$$

Now, choosing arbitrary $n, m \in \{1, 2,..., d\}$ and defining $G_{j,l}^{n,m}$ to be the $(d^2+d)$-vector

$$G_{j,l}^{n,m} \triangleq [0_1 0_2 \cdots 0_{md} (F_{j,l}^{(n)})^T 0_{md+1} \cdots 0_{d^2}]^T \tag{3.20}$$

for $l = 2, 3, 4,...$ and $j = 1, 2,..., l-1$ and noting $F_{j,l}$ is symmetric, we see by Eqs. (3.2) and (3.20) that

$$\left\{\sum_{j=p}^{q} F_{j,l}(X_j^A - X_{j-1}^A)\right\}^{(n,m)} = \sum_{j=p}^{q} (G_{j,l}^{n,m})^T \left[\int_0^j \mathscr{E}(j,s)\, dW_s - \int_0^{j-1} \mathscr{E}(j-1,s)\, dW_s\right] \tag{3.21}$$

for all integers $l \geqslant 2$ and $1 \leqslant p \leqslant q \leqslant l-1$, and for a.a. $\omega$. Hence, by Eq. (3.21), the usual Hilbert space isometry, Eq. (3.20), and Lemma 2, we have that

$$
\begin{aligned}
&\left|\operatorname{var}\left(\left\{\sum_{j=p}^{q} F_{j,l}(X_j^A - X_{j-1}^A)\right\}^{(n,m)}\right)\right| \\
&= \int_{q-1}^{q} |\mathscr{E}^T(q,s)\, G_{q,l}^{n,m}|^2\, ds + \int_0^{q-1} \left| \sum_{j=p \vee \ulcorner s \urcorner}^{q} \mathscr{E}^T(j,s)\, G_{j,l}^{n,m} \right. \\
&\quad \left. - \sum_{j=p \vee \ulcorner s+1 \urcorner}^{q} \mathscr{E}^T(j-1,s)\, G_{j,l}^{n,m} \right|^2 ds \\
&\ll^{q,l} |||F_{q,l}|||^2 + \int_{p-1}^{q-1} |\mathscr{E}^T(\ulcorner s \urcorner, s)\, G_{\ulcorner s \urcorner, l}^{n,m}|^2\, ds \\
&\quad + \int_0^{q-1} \left| \sum_{j=p \vee \ulcorner s+1 \urcorner}^{q} [\mathscr{E}^T(j,s) - \mathscr{E}^T(j-1,s)]\, G_{j,l}^{n,m} \right|^2 ds \\
&\ll^{q,l} \sum_{j=p}^{q} |||F_{j,l}|||^2 + (1+\log(q))^2 \sum_{j=p}^{q} |||F_{j,l}|||^2 \\
&\ll^{q,l} (1+\log(q))^2 \sum_{j=p}^{q} |||F_{j,l}|||^2
\end{aligned} \tag{3.22}
$$

for integers $l \geqslant 2$ and $1 \leqslant p \leqslant q \leqslant l-1$. Moreover, it follows from the Gaussian moment generating function, the fact $\exp(|z|) = \exp(z^+) \vee \exp(z^-)$ for $z \in \mathfrak{R}$ and Eq. (3.22) that there is a constant $c_1 > 0$ such that

$$
\begin{aligned}
&E \exp\left(\theta \left| \sum_{j=p}^{q} \{F_{j,l}(X_j^A - X_{j-1}^A)\}^{(n,m)} \right| \right) \\
&\qquad \leqslant 2 \exp\left(c_1 \theta^2 [1 + \log(n_{r+1}+1)]^2 \sum_{j=p}^{q} |||F_{j,l}|||^2\right)
\end{aligned} \tag{3.23}
$$

for all integers $p, q, r, l$ such that $l \geqslant 2$, $r \in \{0, 1, ..., s_l\}$, $p, q \in I_r$ and $p \leqslant q < l$, and for all $\theta > 0$. Next, by Eq. (3.23) and Theorem 2 we have a $c_2 > 0$ $(= 12c_1)$ such that

$$
\begin{aligned}
&E \exp\left(\theta \max_{\substack{p \in I_r \\ p < l}} \left| \sum_{\substack{j \in I_r \\ j < p}} \{F_{j,l}(X_j^A - X_{j-1}^A)\}^{(n,m)} \right| \right) \\
&\qquad \leqslant 16 \exp\left(c_2 \theta^2 [1 + \log(n_{r+1}+1)]^2 \sum_{\substack{j \in I_r \\ j < l}} |||F_{j,l}|||^2\right)
\end{aligned} \tag{3.24}
$$

for all $l=2, 3, 4,..., r=0, 1,..., s_l$, and $\theta>0$. Consequently, assuming $\sum_{\substack{j\in I_r\\ j<l}} |||F_{j,l}|||^2>0$, letting

$$\theta_{r,l} \triangleq \left(c_2[1+\log(n_{r+1}+1)]^2 \sum_{\substack{j\in I_r\\ j<l}} |||F_{j,l}|||^2\right)^{-1/2}$$

$$\text{for all} \quad l=2, 3, 4,..., r=0, 1,..., s_l \tag{3.25}$$

and fixing an $a>1+\gamma/2$; we have by Eqs. (3.24) and (3.25), monotone convergence, and Remark 5 (a–c) that

$$E\sum_{l=2}^{\infty}\sum_{r=0}^{s_l} \frac{\exp(\theta_{r,l}\max_{\substack{p\in I_r\\ p<l}} |\sum_{\substack{j\in I_r\\ j<p}} \{F_{j,l}(X_j^A-X_{j-1}^A)\}^{(n,m)}|)}{(n_{s_l}+1)^a}$$

$$\leqslant \sum_{l=2}^{\infty} \frac{s_l+1}{(n_{s_l}+1)^a} \leqslant \sum_{l=2}^{\infty} l^{-a+\gamma/2}<\infty \tag{3.26}$$

Moreover, using Eq. (3.10), an argument similar to Eqs. (3.18) and (3.19) as well as Remark 5 (d), we find that

$$\sum_{\substack{j\in I_r\\ j<l}} |||F_{j,l}|||^2 \leqslant^{r,l} (n_{s_l}+1)^{-2\lambda_{\min}} (n_r+1)^{2\lambda_{\min}-1-\gamma/2}$$

$$\text{for all} \quad l=2, 3,..., r=0, 1,..., s_l \tag{3.27}$$

Hence, we have by Eqs. (3.26), (3.25), (3.27), Remark 5 (e), and the arbitrary choice of $m$, $n$ that

$$\max_{\substack{p\in I_l\\ p<l}} \left|\left|\left|\sum_{\substack{j\in I_r\\ j\leqslant p}} F_{j,l}(X_j^A-X_{j-1}^A)\right|\right|\right|$$

$$\leqslant^{r,l} (n_{s_l}+1)^{-\lambda_{\min}} (1+\log(n_{s_l}+1))(n_r+1)^{\lambda_{\min}-1/2-\gamma/4}$$

$$\times[1+\log(n_r+1)] \tag{3.28}$$

for $l=2, 3, 4,..., r=0, 1,..., s_l$, and a.a. $\omega$; and (e) follows by Eqs. (3.19) and (3.28), and simplification. □

The following lemma is used in Theorem 1, Lemma 1 (e), Lemma 3, and Lemma 4 (a).

**Lemma 2.** Suppose $F$ is a $\mathfrak{R}^{(d^2+d)\times(d^2+d)}$-valued class $\mathscr{M}$ function. Then, there is a $c>0$ such that: (i) for integers $p$, $q$ satisfying $1\leqslant p\leqslant q$ and sequences of $\mathfrak{R}^{d^2+d}$-vectors $\{G_j\}_{j=p}^{q}$ one has that

$$\int_0^{q-1}\left[\sum_{j=p\vee\lceil s+1\rceil}^{q}|[\mathscr{E}_F^T(j,s)-\mathscr{E}_F^T(j-1,s)]\,G_j|\right]^2 ds$$
$$\leqslant c[1+\log(q)]^2\cdot\sum_{j=p}^{q}|G_j|^2 \tag{3.29}$$

and (ii) for any integer $q\geqslant 1$ and Lebesgue-measurable function $G\colon(0,q]\to\mathfrak{R}^{d^2+d}$ one has that

$$\int_0^{q-1}\left[\int_{\lceil u\rceil}^{q}|[\mathscr{E}_F^T(s,u)-\mathscr{E}_F^T(\lfloor s\rfloor,u)]\,G_s|\,ds\right]^2 du$$
$$\leqslant c[1+\log(q)]^2\cdot\int_1^q|G_s|^2\,ds \tag{3.30}$$

*Proof.* Inasmuch as the proofs of (i) and (ii) are very similar, we will only prove (i). Letting

$$\beta(j,s)\triangleq|||\mathscr{E}_F^T(j,s)-\mathscr{E}_F^T(j-1,s)|||\qquad\text{for all}\quad s>0, j=\lceil s+1\rceil,\lceil s+2\rceil,\ldots \tag{3.31}$$

$$\gamma(v)\triangleq\sup_{0<s\leqslant q-1}\beta(\lceil s\rceil+v,s)\qquad\text{for all}\quad v=1,2,3,\ldots \tag{3.32}$$

and utilizing a change of variables and Cauchy–Schwarz, we find that the left-hand side of Eq. (3.29) is majorized by

$$\int_0^{q-1}\left[\sum_{j=\lceil s+1\rceil}^{q}1(p,j,q)\,\beta(j,s)\,|G_j|\cdot\sum_{l=\lceil s+1\rceil}^{q}1(p,l,q)\,\beta(l,s)\,|G_l|\right]ds$$
$$\leqslant\sum_{v=1}^{q}\sum_{u=1}^{q}\int_0^{q-1}1(p-\lceil s\rceil,v,q-\lceil s\rceil)\,\beta(\lceil s\rceil+v,s)\,|G_{\lceil s\rceil+v}|$$
$$\times 1(p-\lceil s\rceil,u,q-\lceil s\rceil)\,\beta(\lceil s\rceil+u,s)\,|G_{\lceil s\rceil+u}|\,ds$$
$$\leqslant\left[\sum_{v=1}^{q}\gamma(v)\left(\int_0^{q-1}1(p-\lceil s\rceil,v,q-\lceil s\rceil)\,|G_{\lceil s\rceil+v}|^2\,ds\right)^{1/2}\right]^2$$
$$\leqslant\left[\sum_{v=1}^{q}\gamma(v)\right]^2\sum_{j=p}^{q}|G_j|^2 \tag{3.33}$$

The result follows by the definition of a class $\mathscr{M}$ process. □

The next result, used to establish Lemma 1 (e), is Móricz[16] [Thm. 1] with the definition

$$g(F_{p,k}) \triangleq \begin{cases} \tilde{g}(p+a, k+p+a-1) \\ \qquad \text{if } \; p+k \leqslant b-a \\ \tilde{g}(p+a, b) \\ \qquad \text{if } \; p+k > b-a, \; p \leqslant b-a \\ 0 \quad \text{otherwise} \end{cases} \qquad p=0, 1,...; k=1, 2,...$$

**Theorem 2.** Let $a<b$ be fixed positive integers, $\{\xi_i, \; a \leqslant i \leqslant b\}$ be a sequence of $\mathfrak{R}$-valued random variables and $c>0$ be a constant such that

$$\text{(i)} \quad E\left\{\exp\left(\theta\left|\sum_{i=\alpha}^{\beta} \xi_i\right|\right)\right\} \leqslant c \exp(\theta^2 \tilde{g}(\alpha, \beta))$$
$$\text{for all} \quad a \leqslant \alpha \leqslant \beta \leqslant b, \quad \theta > 0$$

where $\tilde{g}(\alpha, \beta)$ is a nonnegative function satisfying

$$\text{(ii)} \quad \tilde{g}(\alpha, \beta) + \tilde{g}(\beta+1, \gamma) \leqslant \tilde{g}(\alpha, \gamma) \qquad \text{for all} \quad a \leqslant \alpha \leqslant \beta < \gamma \leqslant b$$

Then,

$$E\left\{\exp\left(\theta \max_{a \leqslant \beta \leqslant b}\left|\sum_{i=a}^{\beta} \xi_i\right|\right)\right\} \leqslant 8c \exp(12\theta^2 \tilde{g}(a, b)) \qquad \text{for all} \quad \theta > 0$$

## 4. AN ALMOST SURE INVARIANCE PRINCIPLE FOR $\{u_k = 1, 2, 3,...\}$

The following proposition and Lemma 4 (b) bring forth an almost sure invariance principle for $\{u_k, \; k=1, 2, 3,...\}$.

**Proposition 1.** Under Conditions (C1) and (C2) of Section 2 one has:

$$\text{(i)} \quad |\lfloor t \rfloor u_{\lfloor t+1 \rfloor} - Y_{\lfloor t \rfloor}| \ll' t^{1/2-\eta} \qquad \text{for all} \quad t \geqslant 0 \text{ a.s. } [P]$$

$$\text{(ii)} \quad |u_{\lfloor t+1 \rfloor}| \ll' t^{-1/2}(1+\log(t))^2 \qquad \text{for all} \quad t \geqslant 1 \text{ a.s. } [P]$$

where $\{u_k, \; k=1, 2, 3,...\}$ is defined in Eq. (2.12), $\eta$ is any real constant such that $0<\eta \leqslant \gamma/2$ and $\eta < \lambda_{\min} - 1/2$, and $\{Y_t, \; t \geqslant 0\}$ is as in Theorem 1.

**Remark 7.** Although the constituents of Hypothesis (C1) pertaining to $\{A_l,\ l=1,2,3,...\}$ are no longer germane, we prefer to maintain our initial allotment of assumptions throughout this note and rely on the reader to distinguish what is actually being utilized in each proof.

*Proof of Proposition 1.* (ii) is an immediate consequence of (i) and Lemma 4 (a) so it remains to proof (i). Exerting Hypothesis (C1) (i), we can (and do) fix an $\omega$ such that

$$|Z_t - X_t^{\approx}| \ll^t t^{1/2-\gamma}, \qquad Z_t \triangleq \sum_{j \leqslant \lfloor t \rfloor} z_j \qquad \text{for all} \quad t \geqslant 1 \tag{4.1}$$

Now, we let $D[1,\infty)$ denote the space of right-continuous, left-hand-limit functions from $[1,\infty)$ to $\Re^d$, fix an integer $q \geqslant 1$ and define $\mathscr{H}_q(\phi)$: $D[1,\infty) \to \Re^d$ by

$$\mathscr{H}_q(\phi) \triangleq q\left[\sum_{j=2}^{q} F_{j,q+1}(\phi(j)-\phi(j-1)) + F_{1,q+1}\phi(1)\right] \tag{4.2}$$

where $F_{j,k}$ is defined in Eq. (3.10). Next, recalling Definition Eq. (2.12), we can easily determine that $qu_{q+1} = \mathscr{H}_q(Z) + q\prod_{l=1}^{q}(I-(1/l)A_*)(h_1^1 - h_*)$ so, reasoning similar to Eqs. (3.17)–(3.19), we have by Eq. (3.6) that

$$|qu_{q+1} - \mathscr{H}_q(X^{\approx})| \ll^q q\left|\sum_{j=1}^{q} F_{j,q+1}\tilde{z}_j\right| + q^{1-\lambda_{\min}} \tag{4.3}$$

Moreover, employing an argument similar to Eq. (3.14) and noting that

$$\left|\sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\tilde{z}_j\right| \ll^{i,k} (n_{s_k}+1)^{-\lambda_{\min}}(n_i+1)^{\lambda_{\min}-1/2-\gamma}$$
$$\text{for all} \quad k=2,3,4,...,\ i=0,1,...,s_k \tag{4.4}$$

(by analog to Eqs. (3.16)–(3.19)), we find for a.a. $\omega$ and each $k=2,3,4,...$ that

$$\begin{aligned}(k-1)\left|\sum_{j=1}^{k-1} F_{j,k}\tilde{z}_j\right| &\leqslant (k-1)\sum_{i=0}^{s_k}\left|\sum_{\substack{j\in I_i\\ j<k}} F_{j,k}\tilde{z}_j\right| \\ &\ll^k k^{1-\lambda_{\min}}\left[1+\sum_{i=1}^{s_k} i^{2/\gamma(\lambda_{\min}-1/2-\gamma)}\right] \\ &\ll^k k^{(1-\lambda_{\min})\vee(1/2-\gamma/2)}\end{aligned} \tag{4.5}$$

provided $2/\gamma(\lambda_{\min} - \frac{1}{2} - \gamma) \neq -1$. In this case, we still have a bound in terms of $\eta$ as in the statement of Proposition 1. Now, noting $((s+1)/(q+1))^{A_* - I} = \exp(\log((s+1)/(q+1))\, A_*)((q+1)/(s+1))$ for $0 < s \leqslant q$ and defining

$$\begin{aligned} P_s = P_s^q &\triangleq \prod_{j=\lceil s \rceil + 1}^{q} \left(I - \frac{1}{j} A_*\right) \frac{q}{\lceil s \rceil} \\ Q_s = Q_s^q &\triangleq \exp\left(\log\left(\frac{s+1}{q+1}\right) A_*\right) \frac{q+1}{s+1} \qquad \text{for} \quad 0 < s \leqslant q \end{aligned} \tag{4.6}$$

we attain from Eqs. (4.2) and (4.6) that

$$\mathscr{H}_q(X^{\bar{z}}) - Y_q = \int_0^q P_s^q - Q_s^q \, dX_s^{\bar{z}} \tag{4.7}$$

Moreover, letting $\{\Xi_s, s \geqslant 0\}$ be a $\mathfrak{R}^{(d^2+d)\times(d^2+d)}$-valued process such that each component process $\{\Xi_s^{(m,j)}, s \geqslant 0\}$ $(m, j \in \{1, 2,..., d^2+d\})$ has continuous paths for every $\omega$ and is indistinguishable from $\{\int_0^s \mathscr{E}^{(m,j)}(s, u)\, dW_u^{(j)}, s \geqslant 0\}$, one has, by Condition (C1), and Eq. (3.2), that

$$\{X_t^{\bar{z}}, t \geqslant 0\} = \left\{\int_0^t G\left(dF_s \cdot \sum_{j=1}^{d^2+d} \Xi_s^{(j)}\right) + GW_t, t \geqslant 0\right\}$$
$$\text{for} \quad n = 1, 2,..., d^2+d \tag{4.8}$$

where $G \triangleq [I^d 0^d \cdots 0^d]$ is a $d \times (d^2+d)$-matrix. Defining $R_s^1 \triangleq P_s G$ and $R_s^2 \triangleq Q_s G$ for $0 \leqslant s \leqslant q$, using Protter[18] [Thm. 19, p. 55], and applying Eqs. (4.7) and (4.8); we find that $R^1$ is left continuous with right-hand limits, $R^2$ is continuous and

$$\mathscr{H}_q(X^{\bar{z}}) - Y_q = \int_0^q R_s^1 - R_s^2 \, dW_s + \sum_{j=1}^{d^2+d} \int_0^q (R_s^1 - R_s^2)(dF_s \cdot \Xi_s^{(j)}) \qquad \text{a.s.} \tag{4.9}$$

Next, we recall that each component of $\mathscr{E}^R(F)$ is continuous and of finite variation on compact intervals and utilize Theorem 3 (and its preamble) to obtain for *a.a.* $\omega$ and $i = 1, 2$ that

$$\left\{\sum_{j=1}^{d^2+d} \int_0^q R_s^i(dF_s \cdot \Xi_s^{(j)})\right\}^{(n)} = \sum_{l=1}^{d^2+d} \sum_{j=1}^{d^2+d} \int_0^q \int_u^q R_s^{i(n,l)}(dF_s \cdot \mathscr{E}(s, u))^{(l,j)}\, dW_u^{(j)} \tag{4.10}$$

Hence, by Eqs. (2.4) and (2.5), Stieltjes-type integration by parts, and Eq. (4.6) we obtain

$$\left\{\sum_{j=1}^{d^2+d}\int_0^q R_s^i(dF_s\cdot\Xi_s^{(j)})\right\}^{(n)}$$

$$= \sum_{l,j,m=1}^{d^2+d}\int_0^q\int_u^q R_s^{i(n,l)}((\mathscr{E}^R(F)_u)^{-1})^{(m,j)}\, d\mathscr{E}^R(F)_s^{(l,m)}\; dW_u^{(j)}$$

$$= \sum_{l,j,m=1}^{d^2+d}\int_0^q ((\mathscr{E}^R(F)_u)^{-1})^{(m,j)}\,\{G^{(n,l)}\mathscr{E}^R(F)_q^{(l,m)} - R_u^{i(n,l)}\mathscr{E}^R(F)_u^{(l,m)}$$

$$-\int_u^q \mathscr{E}^R(F)_s^{(l,m)}\, dR_s^{i(n,l)}\}\; dW_u^{(j)}$$

$$= \sum_{l=1}^{d^2+d}\sum_{j=1}^{d^2+d}\int_0^q \{G^{(n,l)}\mathscr{E}^{(l,j)}(q,u) - R_u^{i(n,l)}\delta_{l,j}$$

$$\left.-\int_u^q \mathscr{E}^{(l,j)}(s,u)\, dR_s^{i(n,l)}\right\} dW_u^{(j)}$$

$$=\left\{\int_0^q G\cdot\mathscr{E}(q,u)\, dW_u\right\}^{(n)} - \left\{\int_0^q R_u^i\, dW_u\right\}^{(n)}$$

$$-\left\{\int_0^q\left[\int_u^q dR_s^i\cdot\mathscr{E}(s,u)\right] dW_u\right\}^{(n)} \tag{4.11}$$

$\int_u^q dR_s^1\cdot\mathscr{E}(s,u)$ includes any jump in $R_s^1$ at $u$ but not at $q$. Hence, by Eqs. (4.9) and (4.11), we have

$$\mathscr{H}_q(X^z) - Y_q = -\int_0^q\left[\int_u^q d(R_s^1 - R_s^2)\cdot\mathscr{E}(s,u)\right] dW_u \qquad \text{a.s.} \tag{4.12}$$

and by Eq. (4.12), Lemma 3 (for some $\beta > \eta$) and Borel–Cantelli, we have that

$$|\mathscr{H}_q(X^z) - Y_q| \ll^q q^{1/2-\eta} \qquad \text{a.s.} \tag{4.13}$$

and the proposition follows by Eqs. (4.3), (4.5), and (4.13). □

**Lemma 3.** For any integer $q \geq 1$, class $\mathscr{M}$ function $F$ and $\Re^{d\times(d^2+d)}$-matrix $G$; it follows that

$$\left|\!\left|\!\left|\operatorname{var}\left(\int_0^q\left[\int_u^q d(P_s^q - Q_s^q)\cdot G\mathscr{E}(s,u)\right] dW_u\right)\right|\!\right|\!\right| \ll^q q^{1-2\beta}$$

for some $\beta > \eta$; where $P_s^q$ and $Q_s^q$ are defined in Eq. (4.6), $\mathscr{E}(s, u) = \mathscr{E}_F(s, u)$ is defined in Eq. (2.5), $\{W_u, u \geqslant 0\}$ is a $\mathfrak{R}^{d^2+d}$-valued Brownian motion, and $\eta$ is as in the statement of Proposition 1.

*Proof.* One first fixes an integer $q \geqslant 1$ and a $\beta$ such that $\eta < \beta < (\lambda_{\min} - \frac{1}{2}) \wedge 1$, and notes that by Cauchy–Schwarz one only has to show that

$$\sigma_n^2 \triangleq \sigma_{(n,n)}^2 \left( \int_0^q \left[ \int_u^q d(P_s^q - Q_s^q) \cdot G\mathscr{E}(s, u) \right] dW_u \right) \ll^q q^{1-2\beta} \quad \text{for} \quad n = 1, 2, \ldots, d \tag{4.14}$$

Then, fixing such an $n$ and noting $P_s^q$ and $Q_s^q$ are symmetric, we obtain from Eq. (4.6) that

$$\begin{aligned} \sigma_n^2 \ll^q &\int_0^{q-1} \left| \sum_{l=\lceil u \rceil}^{q-1} \mathscr{E}^T(l, u)\, G^T(P_{l+1}^{(n)} - P_l^{(n)} - Q_{l+1}^{(n)} + Q_l^{(n)}) \right|^2 du \\ &+ \int_0^{q-1} \left| \int_{\lceil u \rceil}^q (\mathscr{E}^T(s, u) - \mathscr{E}^T(\lfloor s \rfloor, u)\, G^T \left( \frac{dQ_s}{ds} \right)^{(n)} ds \right|^2 du \\ &+ \int_0^q \left| \int_u^{\lceil u \rceil} \mathscr{E}^T(s, u)\, G^T \left( \frac{dQ_s}{ds} \right)^{(n)} ds \right|^2 du \end{aligned} \tag{4.15}$$

Moreover, for the first term in Eq. (4.15) we have (when $q > 2$) by Eq. (4.6) and Lemma 2 that

$$\begin{aligned} &\int_0^{q-1} \left| \sum_{l=\lceil u \rceil}^{q-1} \left[ \mathscr{E}^T(\lceil u \rceil, u) + \sum_{j=\lceil u+1 \rceil}^{l} (\mathscr{E}^T(j, u) - \mathscr{E}^T(j-1, u)) \right] \right. \\ &\qquad \left. \times\, G^T(P_{l+1}^{(n)} - P_l^{(n)} - Q_{l+1}^{(n)} + Q_l^{(n)}) \right|^2 du \\ &\leqslant \int_0^{q-1} |||\mathscr{E}^T(\lceil u \rceil, u)|||^2\, |||G^T|||^2\, |Q_{\lceil u \rceil}^{(n)} - P_{\lceil u \rceil}^{(n)}|^2\, du \\ &\quad + \int_0^{q-2} \left[ \sum_{j=\lceil u+1 \rceil}^{q-1} |(\mathscr{E}^T(j, u) - \mathscr{E}^T(j-1, u))\, G^T(Q_j^{(n)} - P_j^{(n)})| \right]^2 du \\ &\ll^q \int_0^{q-1} |Q_{\lceil u \rceil}^{(n)} - P_{\lceil u \rceil}^{(n)}|^2\, du + [1 + \log(q)]^2 \sum_{j=1}^{q-1} |Q_j^{(n)} - P_j^{(n)}|^2 \\ &\ll^q [1 + \log(q)]^2 \sum_{j=1}^{q-1} |||Q_j - P_j|||^2 \end{aligned} \tag{4.16}$$

Now, applying the principle of uniform boundedness, exhibiting $\lambda$ as the eigenvalue of $A_*$ which maximizes each right-hand side, and utilizing Taylor's theorem, we obtain by Eq. (3.17) that

$$\begin{aligned}
|||Q_j^q - P_j^q||| &\ll^{j,q} \frac{q}{j}\left|\prod_{l=j+1}^{q}\left(1-\frac{\lambda}{l}\right) - \exp\left(-\int_{j+1}^{q+1}\frac{\lambda}{s}\,ds\right)\right| \\
&\quad + \left(\frac{q}{j}-\frac{q+1}{j+1}\right)\left(\frac{j+1}{q+1}\right)^{\lambda} \\
&\ll^{j,q} \frac{q}{j}\left\{\left|\prod_{l=j+1}^{q}\left(1-\frac{\lambda}{l}\right) - \prod_{l=j+1}^{q}\exp\left(-\frac{\lambda}{l}\right)\right|\right. \\
&\quad + \left|\exp\left(-\sum_{l=j+1}^{q}\frac{\lambda}{l}\right) - \exp\left(-\int_{j+1}^{q+1}\frac{\lambda}{s}\,ds\right)\right| \\
&\quad \left. + \frac{1}{j+1}\left(\frac{j+1}{q+1}\right)^{\lambda}\right\} \\
&\ll^{j,q} \frac{q}{j}\left\{\sum_{k=j+1}^{q}\exp\left(-\sum_{l=k+1}^{q}\frac{\lambda}{l}\right)\left[\prod_{l=j+1}^{k-1}\left|1-\frac{\lambda}{l}\right|\right]\right. \\
&\quad \times\left|1-\frac{\lambda}{k}-\exp\left(-\frac{\lambda}{k}\right)\right| \\
&\quad \left. + \left(\frac{j+1}{q+1}\right)^{\lambda}\left[\left|\exp\left(\int_{j+1}^{q+1}\frac{\lambda}{s}\,ds - \sum_{l=j+1}^{q}\frac{\lambda}{l}\right)-1\right| + \frac{1}{j+1}\right]\right\} \\
&\ll^{j,q} \left(\frac{j+1}{q+1}\right)^{\lambda_{\min}}\frac{1}{j+1}\frac{q}{j} \qquad \text{for} \quad j = 1, 2, \ldots, q-1 \qquad (4.17)
\end{aligned}$$

Hence, substituting Eq. (4.17) into Eq. (4.16), that

$$\begin{aligned}
&\int_0^{q-1}\left|\sum_{l=\lceil u\rceil}^{q-1}\mathscr{E}^T(l,u)\,G^T(P_{l+1}^{(n)} - P_l^{(n)} - Q_{l+1}^{(n)} + Q_l^{(n)})\right|^2 du \\
&\quad \ll^q [1+\log(q)]^2 \sum_{j=1}^{q-1}\frac{j^{2\lambda_{\min}-4}}{q^{2-2\lambda_{\min}}} \ll^q q^{1-2\beta} \qquad (4.18)
\end{aligned}$$

Next, we note by Lemma 2 and Eq. (4.6) that the second term of Eq. (4.15) can be majorized by

$$\begin{aligned}
&[1+\log(q)]^2\cdot\int_1^q\left|\left|\left|\exp\left(\log\left(\frac{s+1}{q+1}\right)A_*\right)(A_*-I)\frac{q+1}{(s+1)^2}\right|\right|\right|^2 ds \\
&\quad \ll^q [1+\log(q)]^2\,(q+1)^{2-2\lambda_{\min}}\int_1^q (s+1)^{2\lambda_{\min}-4}\,ds \ll^q q^{1-2\beta} \qquad (4.19)
\end{aligned}$$

Finally, we consider the last term in Eq. (4.15) and use the definition of a class $\mathscr{M}$ function to obtain that

$$\int_0^q \left| \int_u^{\lceil u \rceil} \mathscr{E}^T(s,u)\, G^T \left(\frac{dQ_s}{ds}\right)^{(n)} ds \right|^2 du$$

$$\ll^q \int_0^q \left[ \int_u^{\lceil u \rceil} \left(\frac{s+1}{q+1}\right)^{\lambda_{\min}} \frac{q+1}{(s+1)^2}\, ds \right]^2 du$$

$$\ll^q (q+1)^{2-2\lambda_{\min}} \int_0^q (\lceil u \rceil + 1)^{2\lambda_{\min}-4} + (u+1)^{2\lambda_{\min}-4}\, du$$

$$\ll^q q^{1-2\beta} \tag{4.20}$$

and the lemma follows by Eq. (4.15) and (4.18)–(4.20). □

The following stochastic Fubini's theorem was used in Eq. (4.10) of the proof of Proposition 1 with $D(s,u) = 1(0,u,s)\, R_s^{i(n,l)} \mathscr{E}^{(m,j)}(s,u)$, $\mu(ds) = dF_s^{(l,m)}$ and $\beta(s) = R_s^{i(n,l)} \Xi_s^{(m,j)}$ for all possible values of $i$, $j$, $l$, $m$, and $n$. The joint measurability measurability of such a $\beta$ follows from its left continuity. The result follows by a Jordan decomposition on $\mu$, ordinary Fubini theorem, and arguments similar to but simpler than those for Thms. 44 and 45 of Prooter,[18] [pp. 158–159].

**Theorem 3.** Suppose $q > 0$ is a given real number, $D\colon [0,q] \times [0,q] \to \mathfrak{R}$ is a bounded and $\mathscr{B}[0,q] \otimes \mathscr{B}[0,q]$-measurable function, $\mu$ is a finite signed measure on $\mathscr{B}[0,q]$ and $\{W_t, \mathscr{F}_t, t \geqslant 0\}$ is a $\mathfrak{R}$-valued Brownian motion (with $W_0 = 0$) on $(\Omega, \mathscr{F}, P)$. Then, there is a version of $(s,\omega) \to \int_0^q D(s,u)\, dW_u(\omega)$ that is $\mathscr{B}[0,q] \otimes \mathscr{F}$-measurable, and any such jointly-measurable version $(s,\omega) \to \beta(s,\omega)$ is $\mu$-integrable for a.a. $\omega$ and satisfies

$$\int_0^q \beta(s)\, \mu(ds) = \int_0^q \int_0^q D(s,u)\, \mu(ds)\, dW_u \qquad \text{a.s.}$$

where both sides of this equation are $\mathscr{F}$-measurable.

## 5. BOUNDS FOR $\{Y_t, t \geqslant 0\}$

**Lemma 4.** Under Conditions (C1) and (C2) of Section 2, it follows for a.a. $\omega$ $[P]$ that:

(a) $|Y_q| \ll^q q^{1/2}(1+\log(q))^2 \qquad q = 1, 2, 3,\ldots$

(b) $|Y_t - Y_{\lfloor t \rfloor}| \ll^t t^{1/2-\eta} \qquad t \geqslant 0$

where $\eta$ is any real constant such that $0<\eta\leqslant\gamma$ and $\eta<\lambda_{\min}-\frac{1}{2}$, and $\{Y_t, t\geqslant 0\}$ is as in Theorem 1.

*Proof.* Fixing a $q$, one determines analogously to Eqs. (4.9)–(4.11) that for a.a. $\omega$

$$Y_q=\int_0^q G\cdot\mathscr{E}(q,u)\,dW_u-\int_0^q\left[\int_u^q dQ_s^q\cdot G\mathscr{E}(s,u)\right]dW_u \tag{5.1}$$

where $Q_s^q$ is defined in Eq. (4.6). Now, (a) follows by Cauchy–Schwarz, Gaussian tail bounds and Borel-Cantelli if we bound $\sigma^2_{(n,n)}(Y_q)$ appropriately. However, we have by Eq. (5.1), isometry, Lemma 2, the fact $\lambda_{\min}>\frac{1}{2}$ and Eq. (4.6) that

$$\begin{aligned}
\sigma^2_{(n,n)}(Y_q) &\ll^q \int_0^q |\mathscr{E}^T(\lceil u\rceil,u)\,G^T Q^{q(n)}_{\lceil u\rceil}|^2\,du\\
&\quad+\int_0^{q-1}\left|\sum_{j=\lceil u\rceil+1}^{q}[\mathscr{E}^T(j,u)-\mathscr{E}^T(j-1,u)]\,G^T Q_j^{q(n)}\right|^2 du\\
&\quad+\int_0^{q-1}\left|\int_{\lceil u\rceil}^{q}[\mathscr{E}^T(\lfloor s\rfloor,u)-\mathscr{E}^T(s,u)]\,G^T\left(\frac{dQ_s^q}{ds}\right)^{(n)}ds\right|^2 du\\
&\quad+\int_0^{q}\left|\int_u^{\lceil u\rceil}\mathscr{E}^T(s,u)\,G^T\left(\frac{dQ_s^q}{ds}\right)^{(n)}ds\right|^2 du\\
&\ll^q \int_0^q |||Q^q_{\lceil u\rceil}|||^2\,du+[1+\log(q)]^2\left[\sum_{j=2}^{q}|||Q_j^q|||^2\right.\\
&\quad\left.+\int_1^q\left|\left|\left|\frac{dQ_u^q}{du}\right|\right|\right|^2 du\right]+\int_0^q\max_{u\leqslant s\leqslant\lceil u\rceil}\left|\left|\left|\frac{dQ_s^q}{ds}\right|\right|\right|^2 du\\
&\ll^q [1+\log(q)]^2(q+1)^{2-2\lambda_{\min}}\left[\sum_{j=1}^{q}(j+1)^{2\lambda_{\min}-2}\right.\\
&\quad\left.+\int_0^q(u+1)^{2\lambda_{\min}-4}\,du+\sum_{j=1}^{q}(j+1)^{2\lambda_{\min}-4}\right]\\
&\ll^q [1+\log(q)]^2\,q \qquad \text{for}\quad n=1,2,\ldots,d
\end{aligned} \tag{5.2}$$

To prove (b), one fixes an $\omega$ such that Condition (C1) (i) and (3.5) hold, verifies from (C1) (i) that

$$|X_t^z-X^z_{\lfloor t\rfloor}|\ll^t t^{1/2-\gamma}\qquad\text{for all}\quad t>0 \tag{5.3}$$

and deduces from Eq. (5.3) as well as Eq. (3.5) not only that

$$|X^z_{v-1} - X^z_{v((\lfloor t\rfloor+1)/(t+1))-1}|$$

$$\ll^{v,t} (v-1)^{1/2-\gamma} + (1+\log^2(t))\left[\lfloor v\rfloor - \left\lfloor v\frac{\lfloor t\rfloor+1}{t+1}\right\rfloor\right] \tag{5.4}$$

for all $v$, $t$ such that $(t+1)/(\lfloor t\rfloor+1) < v \leqslant t+1$ and $t \geqslant 1$ but also that

$$|X^z_{v-1}| \ll^{v,t} 1 \qquad \text{for all} \quad 1 \leqslant v \leqslant \frac{t+1}{\lfloor t\rfloor+1}, \quad t \geqslant 1 \tag{5.5}$$

Then, by these estimates and Remark 4 one finds for a.a. $\omega$ that

$$\begin{aligned}|Y_t - Y_{\lfloor t\rfloor}| &\leqslant |X^z_t - X^z_{\lfloor t\rfloor}| + \left|\int_{(t+1)/(\lfloor t\rfloor+1)}^{t+1} \frac{A_* - I}{t+1}\left(\frac{v}{t+1}\right)^{A_*-2I}\right.\\ &\quad \left.\times [X^z_{v((\lfloor t\rfloor+1)/(t+1))-1} - X^z_{v-1}]\, dv\right|\\ &\quad + \left|\int_1^{(t+1)/(\lfloor t\rfloor+1)} \frac{A_*-I}{t+1}\left(\frac{v}{t+1}\right)^{A_*-2I} X^z_{v-1}\, dv\right|\\ &\ll^t t^{1/2-\gamma} + (t+1)^{1-\lambda_{\min}}\left\{\int_{(t+1)/(\lfloor t\rfloor+1)}^{t+1} [v^{1/2-\gamma}+1\right.\\ &\quad \left. + \log^2(t)]\, v^{\lambda_{\min}-2}\, dv + 1\right\}\\ &\ll^t t^{1/2-\eta} \qquad \text{for all} \quad t \geqslant 1\end{aligned} \tag{5.6}$$

Moreover, for a.a. $\omega$, one obtains from Remark 4 and Eq. (5.3) that

$$\begin{aligned}|Y_t| &\leqslant |X^z_t| + \left|\int_0^t \frac{A_*-I}{t+1}\left(\frac{v+1}{t+1}\right)^{A_*-2I} X^z_v\, dv\right|\\ &\ll^t t^{1/2-\gamma} + \int_0^t |X^z_v|\, dv \ll^t t^{1/2-\gamma} + t^{3/2-\gamma} \ll^t t^{1/2-\gamma} \qquad \text{for all} \quad 0 < t < 1\end{aligned} \tag{5.7}$$

and (b) follows by Eqs. (5.6) and (5.7). □

## 6. FURTHER EXAMPLES

We give two simple examples of class $\mathscr{M}$ functions and Condition (C1).

**Example 2.** Although our results hold under greater generality, the case of a (continuous) Brownian motion is of paramount importance, due to numerous theorems (see Berkes and Philipp,[3] [Thm. 3], Kuelbs and Philipp,[14] [Thm. 4], Dehling,[5] [Thm. 3], Eberlein,[6,7] [Thms. 3 and 2, respectively] and Philipp,[17] [Thm. 1]) which can be used with the local law of the iterated logarithm to substantiate Conditions (C1) (i) and (ii). When the components of $\{A_k, k=1, 2, 3,...\}$ and $\{z_k, k=1, 2, 3,...\}$ jointly satisfy the conditions of one of these theorems, Condition (C1) would follow provided

$$\left|\sum_{r\leqslant \lfloor t\rfloor} Ez_r\right| \ll^t t^{1/2-\gamma} \quad \text{and} \quad \left|\left|\left|\sum_{r\leqslant \lfloor t\rfloor} (EA_r - A_*)\right|\right|\right| \ll^t t^{1/2-\gamma} \quad \text{for all} \quad t\geqslant 1 \tag{6.1}$$

**Example 3.** We consider

$$X_t = W_t + \int_0^t \frac{I^{d^2+d} - A}{s+1} X_s\, ds \qquad \text{for all} \quad t\geqslant 0 \tag{6.2}$$

with $A - I^{d^2+d}$ symmetric and positive definite and $\{W_t, t\geqslant 0\}$ a Brownian motion, which, by Remark 4 and the previous example, is an essential Gaussian process in the theories of stochastic approximation and adaptive filtering. In this case, one notes that

$$F_s = (I^{d^2+d} - A)\log(s+1)$$

$$\mathscr{E}_F^T(t, s) = \exp\left((I^{d^2+d} - A)\log\left(\frac{t+1}{s+1}\right)\right) \qquad \text{for all} \quad 0\leqslant s\leqslant t \tag{6.3}$$

Now, (i) in the definition of a class $\mathscr{M}$ function is trivial and (ii) and (iii) follow if

$$\sum_{p=1}^{q} \sup_{p-1\leqslant v\leqslant p} \sup_{u\geqslant 0} |||\mathscr{E}_F^T(\lceil u\rceil + v, u) - \mathscr{E}_F^T(\lceil u\rceil + p - 1, u)|||$$

$$\ll^q 1 + \log(q) \tag{6.4}$$

for all $q = 1, 2, 3, \ldots$ However, considering Eq. (6.3) and letting $\lambda$ ($\lambda_m$) denote an arbitrary (the smallest) eigenvalue of $A$, we find by the mean value theorem that

$$\left| \left( \frac{\lceil u \rceil + v + 1}{u+1} \right)^{1-\lambda} - \left( \frac{\lceil u \rceil + p}{u+1} \right)^{1-\lambda} \right|$$
$$\leqslant \frac{1}{u+1} \left( \frac{\lceil u \rceil + p}{u+1} \right)^{-\lambda} \leqslant \frac{(u+1)^{\lambda_m - 1}}{(\lceil u \rceil + p)^{\lambda_m}} \tag{6.5}$$

for all $u \geqslant 0$, $p-1 \leqslant v \leqslant p$ and $p = 1, 2, 3, \ldots$ Hence, by the fact that $\lambda_m > 1$, we find from Eq. (6.5) that

$$\sup_{u \geqslant 0} |||\mathscr{E}_F^T(\lceil u \rceil + v, u) - \mathscr{E}_F^T(\lceil u \rceil + p - 1, u)||| \ll^{p,v} \sup_{u \geqslant 0} \frac{(u+1)^{\lambda_m - 1}}{(u+p)^{\lambda_m}} \ll^{p,v} p^{-1} \tag{6.6}$$

for all $p-1 \leqslant v \leqslant p$, $p = 1, 2, 3, \ldots$, where we have considered the cases $p \geqslant \lceil \lambda_m/(\lambda_m - 1) \rceil$ and $p < \lceil \lambda_m/(\lambda_m - 1) \rceil$ separately. Equation (6.4) follows easily.

## REFERENCES

1. Benveniste, A., Metivier, M., and Priouret, P. (1990). *Adaptive Algorithms and Stochastic Approximations.* Springer, Berlin-Heidelberg-New York.
2. Berger, E. (1986). Asymptotic behavior of a class of stochastic approximation procedures. *Prob. Th. Rel. Fields* **71**, 517–552.
3. Berkes, I., and Philipp, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Prob.* **7**, 29–54.
4. Bradley, R. (1986). Basic properties of strong mixing conditions. In Eberlein, Ernst, and Taqqu, Murad S. (eds.), *Dependence in Probability and Statistics—Survey of Recent Results*, Oberwolfach 1985, Birkhauser, Boston.
5. Dehling, H. (1983). Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. Wahrsch. Verw. Geb.* **63**, 393–432.
6. Eberlein, E. (1983). Strong approximation of very weak Bernoulli processes. *Z. Wahrsch. Verw. Geb.* **62**, 17–37.
7. Eberlein, E. (1986). On strong invariance principles under dependence assumptions. *Ann. Prob.* **14**(1), 260–270.
8. Fabian, V. (1968). On asymptotic normality in stochastic approximation. *Ann. Math. Statist.* **39**, 1327–1332.
9. Farden, D. C. (1981). Stochastic approximation with correlated data. *IEEE Trans. Inform. Theory* **IT-27**(1), 105–113.
10. Gaposhkin, V. F., and Krasulina, T. P. (1974). On the law of the iterated logarithm in stochastic approximation processes. *Theory Prob. Appl.* **19**, 844–850.
11. Heunis, A. J. (1994). Rates of convergence for an adaptive filtering algorithm driven by stationary dependent data. *SIAM J. Control and Optimization* **32**, 116–139.

12. Heyde, C. C. (1974). On martingale limit theory and strong convergence results for stochastic approximation procedures. *Stoch. Proc. Appl.* **2**, 359–370.
13. Kouritzin, M. A. (1994). On the interrelation of almost sure invariance principles for certain stochastic adaptive algorithms and for partial sums of random variables. Technical Report No. 249, Department of Mathematics and Statistics, Carleton University, Ottawa, Canada.
14. Kuelbs, J., and Philipp, W. (1980). Almost sure invariance principles for partial sums of mixing $B$-valued random variables. *Ann. Prob.* **8**(6), 1003–1036.
15. Mark, G. (1982). Loglog-invarianzprinzipien für prozesse der stochastischen approximation. *Mitteilungen Math. Sem. Giessen* **153**, 1–87.
16. Móricz, F. (1976). Probability inequalities of exponential type and laws of the iterated logarithm. *Acta Sci. Math.* **38**, 325–341.
17. Philipp, W. (1986). A note on the almost sure approximation of weakly dependent random variables. *Monatsch. Math.* **102**, 227–236.
18. Protter, P. (1990). *Stochastic Integration and Differential Equations, A New Approach.* Springer, Berlin-Heidelberg-New York.
19. Robbins, H., and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22**, 400–407.
20. Ruppert, D. (1982). Almost sure approximations to the Robbins-Monro and Kiefer-Wolfowitz processes with dependent noise. *Ann. Prob.* **10**, 178–187.
21. Walk, H. (1977). An invariance principle for the Robbins-Monro process in a Hilbert space. *Z. Wahrsch. Verw. Geb.* **39**, 135–150.
22. Walk. H. (1988). Limit behavior of stochastic approximation processes. *Statist. Decisions* **6**, 109–128.
23. Widrow, B., and Stearns, S. D. (1985). *Adaptive Signal Processing.* Prentice-Hall, Englewood Cliffs, New Jersey.