# ASYMPTOTIC THEORY FOR NEAR INTEGRATED PROCESSES DRIVEN BY TEMPERED LINEAR PROCESSES

By

Farzad Sabzikar, Qiying Wang and Peter C. B. Phillips

May 2018

COWLES FOUNDATION DISCUSSION PAPER NO. 2131

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

*Submitted to the Annals of Statistics*

# ASYMPTOTIC THEORY FOR NEAR INTEGRATED PROCESSES DRIVEN BY TEMPERED LINEAR PROCESSES*

By Farzad Sabzikar, Qiying Wang and Peter C. B. Phillips

*Iowa State University, University of Sydney, and Yale University, University of Auckland, University of Southampton, Singapore Management University*

This paper develops an asymptotic theory for near-integrated random processes and some associated regressions when the errors are tempered linear processes. Tempered processes are stationary time series that have a semi-long memory property in the sense that the autocovariogram of the process resembles that of a long memory model for moderate lags but eventually diminishes exponentially fast according to the presence of a decay factor governed by a tempering parameter. When the tempering parameter is sample size dependent, the resulting class of processes admits a wide range of behavior that includes both long memory, semi-long memory, and short memory processes. The paper develops asymptotic theory for such processes and associated regression statistics thereby extending earlier findings that fall within certain subclasses of processes involving near-integrated time series. The limit results relate to tempered fractional processes that include tempered fractional Brownian motion and tempered fractional diffusions. The theory is extended to provide the limiting distribution for autoregressions with such tempered near-integrated time series, thereby enabling analysis of the limit properties of statistics of particular interest in econometrics, such as unit root tests, under more general conditions than existing theory. Some extensions of the theory to the multivariate case are reported.

## 1. Introduction.

Consider a time series that is generated by the model

$$Y(t) = a\, Y(t-1) + X(t), \qquad t = 1, 2, ..., N; \quad Y(0) = 0, \tag{1.1}$$

where $a$ is an unknown parameter and $\{X(j)\}_{j\in\mathbb{Z}}$ is a stationary error process. The observable time series $Y(t)$ in (1.1) is called a near integrated process (or integrated process) when $a$ lies in an $O(N^{-1})$ vicinity of unity (or $a = 1$). Such models have proved useful in applications in many disciplines where observed data show evidence of persistence or randomly wandering behavior. An extensive body of theory now exists concerning the asymptotic properties of data generated by (1.1) and estimators, test statistics and confidence intervals for the autoregressive coefficient $a$. Central to much of this theory is the limit behavior of the ordinary least squares (OLS) estimator

$$\widehat{a} = \frac{\sum_{t=1}^{N} Y(t)Y(t-1)}{\sum_{t=1}^{N} Y^2(t-1)}, \tag{1.2}$$

*Phillips acknowledges research support from the Kelly Fund at the University of Auckland.

*Keywords and phrases:* Tempered linear process, least squares, ARTFIMA, tempered fractional Ornstein-Uhlenbeck process, invariance principle, near integrated process, tempered fractional calculus, semi-long range dependence.

which has been studied under many different assumptions on the structure of the error process $X(t)$.

Assuming $a = a_N := \exp\{c/N\}$, $c \in \mathbb{R}$, in model (1.1), and $\{X(j)\}_{j\in\mathbb{Z}}$ to be weakly dependent errors that satisfy under certain moment and mixing conditions, Phillips [38, Theorem 1] showed that as $N \to \infty$

$$N(\widehat{a}_N - a) \xrightarrow{\mathrm{d}} \Big[\int_0^1 (J_c(s))^2 \,\mathrm{d}s\Big]^{-1}\Big[\int_0^1 J_c dB + \delta\Big] \tag{1.3}$$

$$= \Big[\int_0^1 (J_c(s))^2 \,\mathrm{d}s\Big]^{-1}\Big\{\frac{J_c(1)^2}{2} - c\int_0^1 J_c(s)^2 ds - \frac{\sigma_X^2}{2}\Big\}, \tag{1.4}$$

where $\sigma_X^2 = \mathbb{E}(X(0))^2, \sigma^2 = \sum_{t\in\mathbb{Z}} \mathbb{E}X(0)X(t)$ is the long-run variance of $X(t)$, $\delta = \sum_{t\in\mathbb{N}_+} \mathbb{E}X(0)X(t) = (\sigma^2 - \sigma_X^2)/2$ is a one-sided long run covariance of $X(t)$, and $J_c(r)$ is a linear diffusion (Ornstein-Uhlenbeck) process with Wiener integral

$$J_c(r) = \int_0^r e^{(r-s)c} B(ds), \tag{1.5}$$

based on Brownian motion $B(\cdot)$ with variance $\sigma^2$.

Buchmann and Chan [9] extended this result to the case where the $\{X(j)\}_{j\in\mathbb{Z}}$ are strongly dependent (long memory) errors. In fact, Theorem 2.1 of [9] implies that

$$N^{1\wedge(1+2d)}(\widehat{a}_N - a) \quad \xrightarrow{\mathrm{d}} \quad \frac{1}{\int_0^1 J_{c,d}(s)^2 \,\mathrm{d}s}\begin{cases} \frac{J_{c,d}(1)^2}{2} - c\int_0^1 J_{c,d}(s)^2 ds, & 0 < d < \frac{1}{2}, \\ \frac{J_c(1)^2}{2} - c\int_0^1 J_c(s)^2 ds - \frac{\sigma_X^2}{2}, & d = 0 \\ -\frac{(\Gamma(d+1))^2\sigma_X^2}{2}, & -\frac{1}{2} < d < 0, \end{cases} \tag{1.6}$$

where $J_{c,d}(r)$ is a fractional diffusion process with representation

$$J_{c,d}(r) = \int_0^r e^{(r-s)c} B_d(ds). \tag{1.7}$$

Here $B_d(s)$ is a fractional Brownian motion (fBM) with moving average representation

$$B_d(s) = \frac{1}{\Gamma(d+1)} \int_{\mathbb{R}} \Big[(s-x)_+^d - (-x)_+^d\Big] B(dx).$$

Recently, Sabzikar and Surgailis [45] introduced a class of linear processes called tempered linear processes with semi-long memory properties intermediate between those of long and short memory. A tempered linear process has moving average form

$$X_{d,\lambda}(t) = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda k} b_d(k)\zeta(t-k), \qquad t \in \mathbb{Z} \tag{1.8}$$

in an i.i.d. innovation process $\{\zeta(t)\}$ with $E\zeta(0) = 0$ and $E\zeta^2(0) = 1$, and with coefficients $b_d(k)$ regularly varying at infinity as $k^{d-1}$, viz.,

$$b_d(k) \;\sim\; \frac{c_d}{\Gamma(d)}\, k^{d-1}, \qquad k \to \infty, \qquad c_d \neq 0, \quad d \neq 0, \tag{1.9}$$

where $d \in \mathbb{R}$ is a real number, $d \neq -1, -2, \ldots$, and $\lambda > 0$ is the tempering parameter. A special case of such processes that has been studied in [33, 44, 45] is the two-parameter class of tempered fractionally integrated processes depending only on the parameters $(d, \lambda)$, denoted by ARTFIMA$(0, d, \lambda, 0)$. This class has no autoregressive or moving average component and extends to the tempered process case the well-known class of fractionally integrated autoregressive moving average processes, denoted ARFIMA$(0, d, 0)$. Section B in the Appendix provides definitions and some essential properties of ARTFIMA$(p, d, \lambda, q)$ processes, various specializations, and multivariate extensions. In what follows and given the generality of (1.8), we will mainly focus on ARTFIMA$(0, d, \lambda, 0)$ processes.

When the value of the tempering parameter $\lambda$ is small, an ARTFIMA$(0, d, \lambda, 0)$ process has an autocovariances resembling that of a long memory process out to a large number of lags but eventually decaying exponentially fast. In [19] this behavior was termed *semi-long memory*. Such processes have empirical relevance for modelling time series that are known to display various degrees of long memory with autocovariances that decay slowly at first but ultimately decay much faster, such as the magnitude or certain powers of financial returns (see, for example, [24]).

A specific focus in the present paper is the limit theory associated with the estimator $\widehat{a}_N$ in the regression model (1.1) when the error process follows a tempered linear process given by (1.8) and allowance is made for sample size dependence in the tempering parameter $\lambda = \lambda_N$. This framework extends the usual local to unity asymptotic theory to accommodate a wide class of long memory, intermediate memory, and short memory processes. We consider the following two scenarios:

(a) The parameter $\lambda$ is independent of the sample size $N$; and
(b) The parameter $\lambda = \lambda_N$ depends on $N$ with $\lim_{N\to\infty} N\lambda_N = \lambda^* \in [0, \infty]$.

These cases are analyzed in Section 3 of the paper and the results are summarized as follows. For case (a) the limit distribution of $N(\widehat{a}_N - a)$ follows (1.3) and has the form of a ratio of quadratic functionals of the linear diffusion process (1.5). In case (b) the limit distribution depends on the value of $\lambda^*$. If $\lambda^* \in (0, \infty)$, then the limit distribution modifies (1.6) with the fBM process replaced by a Gaussian stochastic process called tempered fractional Brownian motion of the second kind (TFBM II). But if $\lambda^* = 0$, then (1.6) continues to hold. On the other hand, if $\lambda^* = \infty$, the limit distribution may be written in terms of functionals of standard Brownian motion but these take different forms in the cases $d > 0, d = 0$ and $d < 0$ with $d \neq \mathbb{N}_-$; moreover, except for the case $d = 0$, this limit differ from that of Phillips [38]. The details are given in Theorem 3.3 below.

It is well-known that the process fBM is related to the usual fractional calculus operator. In fact, fractional noise may be interpreted as a fractional integral (derivative) of white noise when $0 < d < \frac{1}{2}$ (respectively, $-\frac{1}{2} < d < 0$) – see [36] for details. A new version of fractional calculus called tempered fractional calculus has been proposed in [14, 44], which usefully relates to tempered fBM. Indeed, working from the Weyl or Riemann-Liouville definition of a fractional operator, a tempered fractional derivative (or integral) replaces the usual power law kernel by a power law kernel scaled by an exponential tempering factor – see [14, 30, 44] for a detailed development. The tempering factor produces a more tractable mathematical object. This tempering factor can be made arbitrarily light and the resulting operator approximates the usual fractional derivative to any desired degree of accuracy over a finite interval. The increment of TFBM II is called tempered fractional Gaussian noise (TFGN II) and it can be shown that TFGN II is the tempered fractional integral (derivative) of the white noise. Readers are referred to [44, 46] for more details on these connections.

Phillips [40] extended the asymptotic results in [38] to the multivariate case by introducing the concept of near-integrated vector processes. Let $\mathbf{Y}(t)$ be a multiple time series that are generated by the model

$$\mathbf{Y}(t) = A\mathbf{Y}(t-1) + \mathbf{X}(t), \tag{1.10}$$

with

$$A = \exp\{N^{-1}C\},$$

where $\{\mathbf{X}(t)\}$ is a weakly stationary sequence of random $m$-vectors that satisfies some mixing conditions, and $C$ is a fixed real $m \times m$ matrix. If $\widehat{A}_N$ is the least squares estimate of $A$ in (1.10), Theorem 4.1 in [40] shows that, as $N \to \infty$,

$$N(\widehat{A}_N - A) \xrightarrow{\mathrm{d}} \Big\{ \int_0^1 dBJ_C' + \Lambda' \Big\} \Big[ \int_0^1 (J_C(s))J_C'(s)\ ds \Big]^{-1}, \tag{1.11}$$

where $J_C(r)$ is a vector diffusion process with stochastic integral representation

$$J_C(r) = \int_0^r e^{(r-s)C} B(ds), \tag{1.12}$$

$B(s)$ is $m$-vector Brownian motion with covariance matrix $\Omega = \sum_{t\in\mathbb{Z}} \mathbb{E}X(0)X(t)'$, the long-run variance matrix of $X(t)$, and $\Lambda = \sum_{t\in\mathbb{N}_+} \mathbb{E}X(0)X(t)'$ is the one-sided long run covariance matrix of $X(t)$. Motivated by (1.11), a result that has proved useful in the study of nonstationary vector autoregressions and power functions for tests of cointegrating rank in econometrics, we consider the regression model (1.10) in the more general setting where the error process follows a strongly tempered linear process. We first establish multivariate invariance principles for the vector of partial sums of $\{\mathbf{X}_{\boldsymbol{d},\boldsymbol{\lambda}}(j)\}$, where $\mathbf{d} = (d_1, \dots, d_m)$ and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)$ – see Theorem 4.1 below. Then, using these results, we develop the limit theory for the sample moments of the tempered near integrated time series (1.10) with additive vector process $\{\mathbf{X}_{\boldsymbol{d},\boldsymbol{\lambda}}(j)\}$ – see Theorem 4.3. Finally, we derive the limit distribution of the ordinary least squares (OLS) regression estimates of the vector time series (1.10) when the errors are strongly tempered – see Theorem 4.2. We emphasize that the approach used to derive asymptotic results for $N(\widehat{A}_N - A)$ in the multivariate case in Section 4 is not simply an extension of the univariate case – see Remark 4.4 below and Phillips [42] for this distinction.

In the above and in what follows, we use the notation $\xrightarrow{\mathrm{d}}$, $\overset{\mathrm{d}}{=}$, and $\xrightarrow{\mathrm{fdd}}$, $\overset{\mathrm{fdd}}{=}$ for weak convergence and equality of distributions, and finite-dimensional weak convergence and equality, respectively. We also write $\Rightarrow$ for weak convergence of random processes in the Skorohod space equipped with $J_1$-topology, see [6], and use the notation $\mathbb{N}_\pm := \{\pm 1, \pm 2, \dots\}, \mathbb{R}_+ := (0,\infty), (x)_\pm := \max(\pm x, 0), x \in \mathbb{R},$ and $\int := \int_{\mathbb{R}}$. $L^p(\mathbb{R})\,(p \ge 1)$ denotes the Banach space of measurable functions $f : \mathbb{R} \to \mathbb{R}$ with finite norm $\|f\|_p = \big(\int |f(x)|^p \mathrm{d}x\big)^{1/p}$. The matrix $\mathrm{diag}(\eta_1, \dots, \eta_m)$ is $m \times m$ diagonal with entries $\eta_1, \dots, \eta_m$. Throughout this paper, all asymptotic results apply as $N \to \infty$.

**2. Tempered fractional processes.** Let $\{B(t)\}_{t\in\mathbb{R}}$ be a two-sided real-valued Brownian motion on the real line, a process with stationary independent increments such that $B(t)$ has a

Gaussian distribution with mean zero and variance $\sigma^2|t|$ for all $t \in \mathbb{R}$, for some $\sigma > 0$. Define an independently scattered Gaussian random measure $B(dx)$ with control measure $m(dx) = \sigma^2 dx$ by setting $B[a,b] = B(b) - B(a)$ for any real numbers $a < b$, and then extending to all Borel sets. Then the stochastic integrals $I(f) := \int_{\mathbb{R}} f(x)B(dx)$ are defined for all functions $f : \mathbb{R} \to \mathbb{R}$ such that $\int f(x)^2 dx < \infty$ as Gaussian random variables with mean zero and covariance $\mathbb{E}[I(f)I(g)] = \sigma^2 \int f(x)g(x)dx$ – see for example [47, Chapter 3].

A fractional Brownian motion (fBM) is a Gaussian stochastic process with the moving average representation

$$B_d(t) = \frac{1}{\Gamma(d+1)} \int_{\mathbb{R}} \left[ (t-x)_+^d - (-x)_+^d \right] B(dx), \tag{2.1}$$

where the memory parameter $d$ satifies $-\frac{1}{2} < d < \frac{1}{2}$. The properties of $B_d(t)$ are explored in detail in [47, Chapter 7]. Meeerschaert and Sabzikar [31] and Sabzikar and Surgailis [46] introduced tempered fractional Brownian motion (TFBM) and tempered fractional Brownian motion of the second kind (TFBM II) respectively. A TFBM is a Gaussian stochastic process with the moving average representation

$$B_{d,\lambda}(t) = \int_{\mathbb{R}} \left[ (t-x)_+^d \ e^{-\lambda(t-x)_+} - (-x)_+^d \ e^{-\lambda(-x)_+} \right] B(dx) \tag{2.2}$$

where $d > -\frac{1}{2}$ and $\lambda > 0$. A TFBM II is a Gaussian stochastic process defined by

$$B^{II}_{d,\lambda}(t) = \int_{\mathbb{R}} h_{d,\lambda}(t,x) B(dx), \tag{2.3}$$

where

$$h_{d,\lambda}(t;x) = (t-x)_+^d \ e^{-\lambda(t-x)_+} - (-x)_+^d \ e^{-\lambda(-x)_+} + \lambda \int_0^t (s-x)_+^d \ e^{-\lambda(s-x)_+} \ ds, \qquad y \in \mathbb{R} \tag{2.4}$$

for $d > -\frac{1}{2}$ and $\lambda > 0$. TFBM and TFBM II reduce to fBM when $\lambda = 0$ and $-\frac{1}{2} < d < \frac{1}{2}$. In this paper, since our results relate closely to TFBM II, it will be useful to summarize the basic properties of $B^{II}_{H,\lambda}(t)$. Readers are referred to [46] for the details.

**Proposition 2.1** (i) TFBM II $B^{II}_{H,\lambda}$ in (2.3) has stationary increments, such that

$$\left\{ B^{II}_{d,\lambda}(ct) \right\}_{t \in \mathbb{R}} \overset{\text{fdd}}{=} \left\{ c^{d+\frac{1}{2}} B^{II}_{d,\lambda t}(t) \right\}_{t \in \mathbb{R}} \tag{2.5}$$

for any scale factor $c > 0$ and is not a self-similar process.

(ii) TFBM II $B^{II}_{d,\lambda}$ in (2.3) has a.s. continuous paths.

(iii) For $d > 0$, the covariance function of TFBM II $B^{II}_{d,\lambda}$ is given by

$$\mathbb{E} B^{II}_{d,\lambda}(t) B^{II}_{d,\lambda}(s) = C(H,\lambda) \int_0^t \int_0^s |u-v|^{d-\frac{1}{2}} K_{d-\frac{1}{2}}(\lambda|u-v|) dv \ du, \tag{2.6}$$

where $C(d,\lambda) = \frac{2}{\sqrt{\pi}\Gamma(d)(2\lambda)^{d-\frac{1}{2}}}$, $d > 0$, and $\lambda > 0$. Here $K_\nu(x)$ is the modified Bessel function of the second kind (see [1, Chapter 9]).

**Remark 2.2** For $d > \frac{1}{2}$ the integrand in (2.6), viz.,

$$\frac{1}{\sqrt{\pi}\Gamma(d)(2\lambda)^{d-\frac{1}{2}}}|u-v|^{d-\frac{1}{2}}K_{d-\frac{1}{2}}(\lambda|u-v|) \tag{2.7}$$

is the Matérn covariance function (in one dimension) with shape parameter $\nu = d - \frac{1}{2} > 0$, scale parameter $\lambda > 0$, and variance parameter 1, see e.g. ([7], (1.1)). Note that the integral in (2.6) diverges when $-\frac{1}{2} < d < 0$. A more complex representation of the covariance function of $B^{II}_{d,\lambda}$ is available for the case $-\frac{1}{2} < d < 0$, but it is not needed in the present paper.

Next, we define the following stochastic process that plays an important role in the limit distribution theory.

**Definition 2.3** *A tempered fractional Ornstein-Uhlenbeck (OU) process of the second kind (TFOU II) is defined as*

$$J^{II}_{c,d,\lambda}(r) = \int_0^r e^{(r-s)c} dB^{II}_{d,\lambda}(s), \tag{2.8}$$

*where* $\{B^{II}_{d,\lambda}(s)\}_{s\in\mathbb{R}}$ *is the TFBM II given by* (2.3) .

**Lemma 2.4** *Let* $J^{II}_{c,d,\lambda}$ *be the TFOU II given by* (2.8). *Then* $J^{II}_{c,d,\lambda}$ *is a Gaussian stochastic process with zero mean and finite variance.*

**Remark 2.5** It can be shown that TFOU II is the unique solution of the following Langevin equation driven by a TFBM II process

$$dJ^{II}_{c,d,\lambda}(r) = cJ^{II}_{c,d,\lambda}(r)dr + \theta dB^{II}_{d,\lambda}(r). \tag{2.9}$$

under the initial condition $\xi^{II}_{d,\lambda} = \theta \int_{-\infty}^0 e^{cr} dB^{II}_{d,\lambda}(r)$.

We close this section with a discussion of the tempered fractionally integrated process that is a special case of tempered linear process given by (1.8). An ARTFIMA$(0, d, \lambda, 0)$ class of tempered fractionally integrated processes, generalizing the well-known ARFIMA$(0, d, 0)$ class, is defined by

$$X_{d,\lambda}(t) = (1 - \mathrm{e}^{-\lambda}B)^{-d}\zeta(t) = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda k}\omega_{-d}(k)\zeta(t-k), \qquad t \in \mathbb{Z} \tag{2.10}$$

with coefficients given by power expansion $(1 - \mathrm{e}^{-\lambda}z)^{-d} = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda k}\omega_{-d}(k)z^k, |z| < 1$, where $\omega_{-d}(k) := \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}$ for $d \in \mathbb{R}\backslash\mathbb{N}_-$ and $Bx(t) = x(t-1)$ is the backward shift. Due to the presence of the exponential tempering factor $\mathrm{e}^{-\lambda k}$ the series in (1.8) and (2.10) converges absolutely a.s. and in $L_p$ under general assumptions on the innovations and thereby defines a strictly stationary process.

**Remark 2.6** (i) Time series in the ARTFIMA$(0, d, \lambda, 0)$ class have covariance function

$$\gamma_{d,\lambda}(k) = \mathbb{E}X_{0,d,\lambda,0}(0)X_{0,d,\lambda,0}(k) = \frac{\mathrm{e}^{-\lambda k}\Gamma(d+k)}{\Gamma(d)\Gamma(k+1)}\, {}_2F_1(d, k+d; k+1; \mathrm{e}^{-2\lambda}), \tag{2.11}$$

where ${}_2F_1(a, b; c; z)$ is the Gauss hypergeometric function (see e.g. [22]). Moreover,

$$\sum_{k\in\mathbb{Z}} \left|\gamma_{d,\lambda}(k)\right| < \infty, \qquad \sum_{k\in\mathbb{Z}} \gamma_{d,\lambda}(k) = (1 - \mathrm{e}^{-\lambda})^{-2d} \tag{2.12}$$

and

$$\gamma_{d,\lambda}(k) \sim A k^{d-1} \mathrm{e}^{-\lambda k}, \quad k \to \infty, \quad \text{where} \quad A = (1-\mathrm{e}^{-2\lambda})^{-d}\Gamma(d)^{-1}. \tag{2.13}$$

(ii) From (2.13) it is evident that for small values of $\lambda$ the covariance function of the ARTFIMA model may resemble the covariance function of a long memory model out to a large number of lags but eventually decays exponentially fast. [19] termed such behavior 'semi long-memory' and noted that models generating such time series may have empirical relevance for capturing certain long-run features of financial returns ([24]).
(iii) The ARTFIMA$(0, d, \lambda, 0)$ class can be extended to ARTFIMA$(p, d, \lambda, q)$ in two different ways, as explained in Appendix B. However, the present paper mainly focuses on the ARTFIMA$(0, d, \lambda, 0)$ class.

**3. Near integrated processes with ARTFIMA innovations.** This section develops asymptotic theory for near-integrated processes with ARTFIMA innovations and for autoregression with such processes. We first study the asymptotic theory for the sample moments of data generated by the autoregression (1.1) when $a = \exp\{c/N\}$ and $X = X_{d,\lambda}$ is ARTFIMA$(0, d, \lambda_N, 0)$ given by (2.10) with $\mathbb{E}\zeta(0) = 0$ and $\mathbb{E}\zeta^2(0) = 1$. These results are employed to obtain the limit distribution of the fitted autoregressive coefficient $\hat{a}_N$, which depends on the TFOU II process – see Theorem 3.3 below. In the following, to simplify notation we write $J_c = J^{II}_{c,0,0}$ and $J_{c,d} = J^{II}_{c,d,0}$ where $J^{II}_{c,d,\lambda}$ is the TFOU II process given by (2.8).

**Lemma 3.1** *(i) Let $\lambda^* = \infty$, $d \in \mathbb{R} \setminus \mathbb{N}_-$. Then*

$$N^{-1/2}\lambda_N^d Y[Ns] \Rightarrow J_c(s)$$

*on $D[0,1]$ and*

$$N^{-2}\lambda_N^{2d} \sum_{t=1}^{N} Y^2(t-1) \xrightarrow{\mathrm{d}} \int_0^1 J_c(s)^2\, ds.$$

*(ii) Let $\lambda^* = 0$ and $-\frac{1}{2} < d < \frac{1}{2}$. Then*

$$N^{-(d+1/2)} Y[Ns] \Rightarrow \Gamma(d+1)^{-1} J_{c,d}(s)$$

*on $D[0,1]$ and*

$$N^{-(2d+2)} \sum_{t=1}^{N} Y^2(t-1) \xrightarrow{\mathrm{d}} \Gamma(d+1)^{-2} \int_0^1 J_{c,d}(r)^2\, dr.$$

*(iii) Let $\lambda^* \in (0,\infty)$ and $d > -\frac{1}{2}$. Then*

$$N^{-(d+\frac{1}{2})} Y[Ns] \Rightarrow \Gamma(d+1)^{-1} J^{II}_{c,d,\lambda^*}(s)$$

*on $D[0,1]$ and*

$$N^{-(2d+2)} \sum_{t=1}^{N} Y^2(t-1) \xrightarrow{\mathrm{d}} \Gamma(d+1)^{-2} \int_0^1 J^{II}_{c,d,\lambda^*}(r)^2 dr.$$

The following proposition is used in deriving the limit distributions in Theorem 3.3, c.f., [45, Proposition 5.1].

**Proposition 3.2** *Let $X_{d,\lambda_N} \equiv X_{0,d,\lambda_N,0}$ be an ARTFIMA$(0,d,\lambda_N,0)$ process in (2.10) with i.i.d. innovations $\{\zeta(t)\}, \mathbb{E}\zeta(0)=0, \mathbb{E}\zeta^2(0)=1$, fractional parameter $d \in \mathbb{R}\setminus\mathbb{N}_-$ and tempering parameter $\lambda_N \to 0$. Moreover, let $\mathbb{E}|\zeta(0)|^p < \infty$, for some $p>2$. Then*

$$\frac{1}{N}\sum_{t=1}^{N} X^2_{d,\lambda_N}(t) \xrightarrow{p} \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}, \qquad d<1/2, \tag{3.1}$$

$$\frac{\lambda_N^{2d-1}}{N}\sum_{t=1}^{N} X^2_{d,\lambda_N}(t) \xrightarrow{p} \frac{\Gamma(d-1/2)}{2\sqrt{\pi}\,\Gamma(d)}, \qquad d>1/2, \tag{3.2}$$

$$\frac{1}{N|\log\lambda_N|}\sum_{t=1}^{N} X^2_{d,\lambda_N}(t) \xrightarrow{p} \frac{1}{\pi}, \qquad d=1/2. \tag{3.3}$$

**Theorem 3.3** *Consider the AR(1) model*

$$Y(t) = aY(t-1) + X_{d,\lambda}(t),$$

*where $a=a_N=\exp\{c/N\}$ and the error process $\{X_{d,\lambda}(j)\}_{j\in\mathbb{Z}}$ is given by (2.10). Assume $\{\zeta(t), t\in\mathbb{Z}\}$ are i.i.d. innovations $\mathbb{E}\zeta(0)=0, \mathbb{E}\zeta^2(0)=1, \mathbb{E}|\zeta(0)|^p<\infty$, for some $p > 2\vee 1/(d+1/2)$, fractional parameter $d\in\mathbb{R}\setminus\mathbb{N}_-$, and tempering parameter $\lambda_N>0$ satisfying $\lim_{N\to\infty} N\lambda_N = \lambda^* \in [0,\infty]$. Let $\widehat{a}_N$ be the OLS estimator of the parameter $a$ given by (1.2).*

*(i)* (Strongly tempered errors.) *Let $\lambda^*=\infty$, $d\in\mathbb{R}\setminus\mathbb{N}_-$. Then*

$$\min(1,\lambda_N^{-2d})N(\widehat{a}_N - a) \xrightarrow{\mathrm{d}} \frac{1}{2\int_0^1 J_c(s)^2\,\mathrm{ds}}\begin{cases} J_c(1)^2 - 2c\int_0^1 (J_c(s))^2 ds, & d>0,\\ J_c(1)^2 - 2c\int_0^1 J_c(s)^2 ds - 1, & d=0,\\ -\frac{\Gamma(1-2d)}{\Gamma(1-d)^2}, & -\frac{1}{2}<d<0,\end{cases}$$

*where $J_c(s) = J^{II}_{c,0,0}(s)$ is given by (2.8).*

*(ii)* (Weakly tempered errors.) *Let $\lambda^*=0$ and $-\frac{1}{2}<d<\frac{1}{2}$. Then*

$$N^{1\wedge(1+2d)}(\hat{a}_N - a) \xrightarrow{\mathrm{d}} \frac{1}{2\int_0^1 (J_{c,d}(s))^2\,\mathrm{ds}}\begin{cases} (J_{c,d}(1))^2 - 2c\int_0^1 (J_{c,d}(s))^2 ds, & 0<d<\frac{1}{2},\\ (J_{c,d}(1))^2 - 2c\int_0^1 (J_{c,d}(s))^2 ds - 1, & d=0\\ -\frac{\Gamma(d+1)^2\Gamma(1-2d)}{\Gamma(1-d)^2}, & -\frac{1}{2}<d<0,\end{cases}$$

*where $J_{c,d} = J^{II}_{c,d,0}(s)$ is given by (2.8).*

*(iii)* (Moderately tempered errors.) *Let $0<\lambda^*<\infty$ and $d>-\frac{1}{2}$. Then*

$$N^{1\wedge(1+2d)}(\hat{a}_N - a) \xrightarrow{\mathrm{d}} \frac{1}{2\int_0^1 (J^{II}_{c,d,\lambda^*}(s))^2\,\mathrm{ds}}\begin{cases} (J^{II}_{c,d,\lambda^*}(1))^2 - 2c\int_0^1 (J^{II}_{c,d,\lambda^*}(s))^2 ds, & d>0,\\ (J^{II}_{c,d,\lambda^*}(1))^2 - 2c\int_0^1 (J^{II}_{c,d,\lambda^*}(s))^2 ds - 1, & d=0\\ -\frac{\Gamma(d+1)^2\Gamma(1-2d)}{\Gamma(1-d)^2}, & -\frac{1}{2}<d<0,\end{cases}$$

*where $J^{II}_{c,d,\lambda^*}(s)$ is given by (2.8).*

**4. Near integrated multiple time series with strongly tempered innovations.** In this section, we extend Theorem 3.3 to the multivariate case when the errors are strongly tempered. We first establish a multivariate generalization of the invariance principles for tempered fractionally integrated processes due to Sabzikar and Surgailis [45] – see Theorem 4.1 below. We then obtain limit theory for the sample moments of a near integrated vector process with strongly tempered errors.

Let $\zeta(t) = (\zeta_1(t), ..., \zeta_m(t))', t \in \mathbb{Z}$, be a time series of iid random vectors with $\mathbb{E}\zeta(t) = 0$ and covariance matrix $\Omega$. Define a random $m$-vector of tempered linear processes

$$\mathbf{X}_{\boldsymbol{d},\boldsymbol{\lambda}}(t) = \Big(X_{d_1,\lambda_1}(t), \ldots, X_{d_m,\lambda_m}(t)\Big)' \tag{4.1}$$

such that, as in (1.8), $X_{d_i,\lambda_i}(t)$ is given by,

$$X_{d_i,\lambda_i}(t) = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda_i k} b_{d_i}(k)\zeta_i(t-k), \quad b_{d_i}(k) \sim \frac{c_{d_i}}{\Gamma(d_i)} k^{d_i-1}.$$

Define the vector partial sums

$$\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t) := \sum_{k=1}^{[Nt]} \mathbf{X}_{\mathbf{d},\boldsymbol{\lambda}}(k), \quad t \in [0,1]. \tag{4.2}$$

Throughout this section, for all $i = 1, \ldots, m$, we assume $d_i > 0$, the tempering parameter $\lambda_i \equiv \lambda_{i,N} \to 0$ as $N \to \infty$ and

$$\lim_{N\to\infty} N\lambda_{i,N} = \infty. \tag{4.3}$$

Following [45], $X_{d_i,\lambda_N}$ is called strongly tempered. We further assume $c_{d_i} = 1, i = 1, ..., m$, for convenience of presentation.

Our first result is the weak convergence of $\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t)$, extending [45] from univariate to multivariate settings. Unlike [45], only the second moment is required to establish the limit theory in this case. Let $D_N = \mathrm{diag}(N^{-\frac{1}{2}}\lambda_1^{d_1}, \cdots, N^{-\frac{1}{2}}\lambda_m^{d_m})$ and $\mathbf{B}(t) = (B_1(t), \ldots, B_m(t))'$ be $m$-dimensional Brownian motion with covariance matrix $\Omega$.

**Theorem 4.1** *We have*

$$D_N\, \mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t) \quad \Rightarrow \quad \mathbf{B}(t), \tag{4.4}$$

*on* $D_{R^m}[0,1]$.

For the multiple times series $\mathbf{Y}(t) = (Y_1(t), ..., Y_m(t))', t \geq 1$, generated by

$$\mathbf{Y}(t) = A\mathbf{Y}(t-1) + \mathbf{X}_{\boldsymbol{d},\lambda}(t), \quad \mathbf{Y}(0) = 0,$$

where $A = \mathrm{diag}(\exp\{c_1/N\}, \ldots, \exp\{c_m/N\})$, as in [39], the coefficient matrix $A$ can be estimated by vector autoregression giving

$$\hat{A}_N = \Big[\sum_{t=1}^{N} \mathbf{Y}(t)\, \mathbf{Y}(t-1)'\Big] \Big[\sum_{t=1}^{N} \mathbf{Y}(t-1)\, \mathbf{Y}(t-1)'\Big]^{-1}.$$

The next theorem gives a partial multivariate generalization of Theorem 3.3.

**Theorem 4.2** *Suppose that $\mathbb{E}||\zeta(0)||^4 < \infty$ and $\lambda_{iN}/\lambda_{jN} \to \eta_{ij} \in [0,\infty]$ as $N \to \infty$. Then,*

$$ND_N(\hat{A}_N - A)D_N^{-1} \xrightarrow{\mathrm{d}} \Big[\int_0^1 d\mathbf{B}(s)J_C(s)' + \Delta\Big]\Big[\int_0^1 J_C(s)J_C(s)'ds\Big]^{-1},$$

*where $J_C(s) = \big(Q_{1s}, ..., Q_{ms}\big)'$ with $Q_{js} = \int_0^s e^{(s-u)c}dB_j(u)$ and $\Delta = (\Delta_{ij})_{m\times m}$ with*

$$\Delta_{ij} = \begin{cases} \frac{1}{2}E\zeta_i^2(0), & \text{if } i = j, \\ \frac{E\big[\zeta_i(0)\zeta_j(0)\big]}{\Gamma(d_1)\Gamma(d_2)} \int_0^\infty x^{d_j-1}e^{-x}dx \int_{\eta_{ij}x}^\infty y^{d_i-1}e^{-y}dy, & \text{if } i \neq j. \end{cases}$$

Note that $D_{ij} = \mathbb{E}\big[\zeta_i(0)\zeta_j(0)\big]$ if $i \neq j$ and $\eta_{ij} = 0$, and $D_{ij} = 0$ if $i \neq j$ and $\eta_{ij} = \infty$.

Let $\widehat{\mathbf{Y}}(t) = D_N\,\mathbf{Y}(t)$. Note that $\frac{1}{N}\sum_{t=1}^N \widehat{\mathbf{Y}}(t-1)\,\widehat{\mathbf{Y}}(t-1)' = \int_0^1 \widehat{\mathbf{Y}}([Ns])\widehat{\mathbf{Y}}([Ns])'ds$ and

$$ND_N(\hat{A}_N - A)D_N^{-1} = \Big[\sum_{t=1}^N D_N\mathbf{X}_{\boldsymbol{d},\lambda}(t)\,\widehat{\mathbf{Y}}(t-1)'\Big]\Big[\frac{1}{N}\sum_{t=1}^N \widehat{\mathbf{Y}}(t-1)\,\widehat{\mathbf{Y}}(t-1)'\Big]^{-1}.$$

Theorem 4.2 follows directly from the continuous mapping theorem and the following theorem.

**Theorem 4.3** *Suppose that $\mathbb{E}||\zeta(0)||^4 < \infty$ and $\lambda_{iN}/\lambda_{jN} \to \eta_{ij} \in [0,\infty]$ as $N \to \infty$. We have*

$$\big(\widehat{\mathbf{Y}}([Ns]),\ \sum_{t=1}^N D_N\mathbf{X}_{\boldsymbol{d},\lambda}(t)\,\widehat{\mathbf{Y}}(t-1)'\big) \Rightarrow \big(J_C(s), \int_0^1 d\mathbf{B}(s)J_C(s)' + \Delta\big), \tag{4.5}$$

*on $D_{R^m}[0,1] \times R^{m\times m}$.*

**Remark 4.4** *In the proof of Theorem 4.3, we need to investigate asymptotics for components of the form $\frac{\lambda_i^{d_i}\lambda_j^{d_j}}{N}\sum_{t=1}^N X_{d_i,\lambda_i}(t)Y_j(t-1)$, which seems difficult without assuming $\lambda_{i,N}N \to \infty$ when $i \neq j$. As a consequence, we have been unable to establish Theorem 4.2 in the weakly and moderately tempered errors cases in the present work. We plan to investigate this case in later research.*

## 5. Proofs.

*Proof of Lemma 2.4* First we note that

$$J_{c,d,\lambda}^{II}(r) = \int_0^r e^{(r-s)c}B_{d,\lambda}^{II}(ds) = \int_{\mathbb{R}} e^{cx}\mathbf{1}_{\{0<x<r\}}B_{d,\lambda}^{II}(dx) = \int_{\mathbb{R}} \Big(\mathbb{I}_-^{d,\lambda}f\Big)(y)B(dy),$$

where $f(x) = e^{cx}\mathbf{1}_{\{0<x<r\}}$. Therefore, using Definition A.4, $J_{c,d,\lambda}^{II}$ is well-defined if we show that $f \in \mathcal{A}_1$. That is (i) $f \in L^2(\mathbb{R})$ and (ii) $\int_{\mathbb{R}}\big|\big(\mathbb{I}_-^{d,\lambda}f\big)(y)\big|^2dy < \infty$. The first condition (i) obviously holds. For the second one, use the Plancherel Theorem to see that $\|\mathbb{I}_-^{d,\lambda}f\|_2^2 = \|\mathcal{F}\big[\mathbb{I}_-^{d,\lambda}f\big]\|_2^2 < \infty$ for all $d > -\frac{1}{2}$, where $\mathcal{F}[f](f) = \hat{f}(k) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-ikx}f(x)dx$ is the Fourier transform of function $f$. In fact, we have

$$\begin{aligned}\|\mathcal{F}\big[\mathbb{I}_-^{d,\lambda}f\big]\|_2^2 &= \int_{\mathbb{R}} |\hat{f}(k)|^2(\lambda^2+k^2)^{-2d} \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{1-2e^{cr}\cos kr + e^{2cr}}{2\pi(c^2+k^2)}(\lambda^2+k^2)^{-d}dk\end{aligned}$$

which is finite if $d > -\frac{1}{2}$.

*Proof of Lemma 3.1.* The idea of the proof is to use the continuous mapping theorem and Theorem 4.3 of Sabzikar and Surgaillis (2017), i.e., on $D[0,1]$,

$$\begin{aligned} S_N^{d,\lambda_N}([Ns]) &:= N^{-\frac{1}{2}}\lambda_N^d \sum_{j=1}^{[Ns]} X_{d,\lambda_N}(j) \\ &\Rightarrow \begin{cases} B^{II}_{0,0}(s), & \text{under (i)} \\ \frac{1}{\Gamma(d+1)} B^{II}_{d,0}(s), & \text{under (ii)} \\ \frac{1}{\Gamma(d+1)} B^{II}_{d,\lambda^*}(s), & \text{under (iii)} \end{cases} . \end{aligned} \tag{5.1}$$

We only prove (i). The other derivations are similar and the details are omitted. The second part of (i) is simple. In fact, by noting that

$$N^{-2}\lambda_N^{2d}\sum_{t=1}^{N} Y^2(t-1) = \int_0^1 \big(\frac{\lambda_N^d}{\sqrt{N}} Y([Ns])\big)^2 ds + o_P(1),$$

the result follows from the first part of (i), i.e., $\frac{\lambda_N^d}{\sqrt{N}}Y([Ns]) \Rightarrow J_c(s)$ and the continuous mapping theorem.

To prove $\frac{\lambda_N^d}{\sqrt{N}}Y([Ns]) \Rightarrow J_c(s)$, it suffices to show

(a) the tightness of $\frac{\lambda_N^d}{\sqrt{N}}Y([Ns])$, and
(b) finite dimensional convergence of $\frac{\lambda_N^d}{\sqrt{N}}Y([Ns])$.

Let $S_N^{d,\lambda_N}(0) = 0$. For any $0 \le m < n$, we have

$$\begin{aligned} &\frac{\lambda_N^d}{\sqrt{N}}(Y(n) - Y(m)) \\ = &\sum_{k=m+1}^{n} e^{(n-k)c/N}(S_N^{d,\lambda_N}(k) - S_N^{d,\lambda_N}(k-1)) + \big[e^{(n-m)c/N} - 1\big]\frac{\lambda_N^d}{\sqrt{N}}Y(m) \\ = &S_N^{d,\lambda_N}(n) - e^{(n-m)c/N}S_N^{d,\lambda_N}(m) + (e^{c/N}-1)\sum_{k=m+1}^{n} e^{(m-k)c/N}S_N^{d,\lambda_N}(k) \\ &+\big[e^{(n-m)c/N} - 1\big]\frac{\lambda_N^d}{\sqrt{N}}Y(m). \end{aligned} \tag{5.2}$$

This yields (by letting $m = 0$)

$$\frac{\lambda_N^d}{\sqrt{N}}\max_{1\le k\le N}|Y(k)| \le \max_{1\le k\le N}|S_N^{d,\lambda_N}(k)|\,\big[1 + N(e^{c/N}-1)\big] \le C\max_{1\le k\le N}|S_N^{d,\lambda_N}(k)|,$$

and, for any $0 \le s < t \le 1$,

$$\frac{\lambda_N^d}{\sqrt{N}}|Y([Nt]) - Y([Ns])| \;\;\le\;\; |S_N^{d,\lambda_N}([Nt]) - S_N^{d,\lambda_N}([Ns])| + C\,(t-s)\max_{1\le k\le N}|S_N^{d,\lambda_N}(k)|.$$

As a consequence, we have proved the tightness of $\frac{\lambda_N^d}{\sqrt{N}}Y([Nt])$, $0 \le t \le 1$, since $S_N^{d,\lambda_N}([Nt]), 0 \le t \le 1$, is tight due to (5.1).

We next prove the finite dimensional convergence of $\frac{\lambda_N^d}{\sqrt{N}}Y([Ns])$. Without loss of generality, we only show $\frac{\lambda_N^d}{\sqrt{N}}Y(N) \xrightarrow{\text{d}} J_c(1)$, since the general situation is a natural application of the Cramér-Wold device. Note that

$$\begin{aligned}
\frac{\lambda_N^d}{\sqrt{N}}Y(N) &= \sum_{k=1}^{N} e^{(N-k)c/N}(S_N^{d,\lambda_N}(k) - S_N^{d,\lambda_N}(k-1)) \\
&= S_N^{d,\lambda_N}(N) + N(e^{c/N}-1)\int_0^{(N-1)/N} e^{(N-1-[Ns])c/N} S_N^{d,\lambda_N}([Ns])ds.
\end{aligned}$$

It follows from $N(e^{c/N}-1) \to c$, $e^{(N-1-[Ns])c/N} \to e^{(1-s)c}$ uniformly in $s \in [0,1]$ and (5.1) that

$$\frac{\lambda_N^d}{\sqrt{N}}Y(N) \xrightarrow{\text{d}} B_{0,0}^{II}(1) + c\int_0^1 e^{(1-s)c}B_{0,0}^{II}(s)ds = J_c(1),$$

as required. The proof of Lemma 3.1 is complete. □

*Proof of Theorem 3.3.* The idea is to use Lemma 3.1 and the continuous mapping theorem. Since all derivations are similar, we only prove part (i) with $d > 0$ in detail.

When $d > 0$, $\min(1, \lambda_N^{-2d}) = 1$ and then

$$\min(1, \lambda_N^{-2d})N(\hat{a_N} - a) = \frac{N^{-1}\lambda_N^{2d}\sum_{t=1}^{N} Y(t-1)X_{d,\lambda_N}(t)}{N^{-2}\lambda_N^{2d}\sum_{t=1}^{N} Y^2(t-1)} \tag{5.3}$$

Note that

$$Y^2(N) = \left(e^{2c/N}-1\right)\sum_{t=1}^{N} Y^2(t-1) + \sum_{t=1}^{N} X_{d,\lambda_N}(t)^2 + 2e^{c/N}\sum_{t=1}^{N} Y(t-1)X_{d,\lambda_N}(t).$$

We may write

$$\begin{aligned}
&N^{-1}\lambda_N^{2d}\sum_{t=1}^{N} Y(t-1)X_{d,\lambda_N}(t) \\
&= \frac{1}{2}e^{-c/N}N^{-1}\lambda_N^{2d}Y^2(N) - \frac{1}{2}e^{-c/N}N(e^{2c/N}-1)N^{-2}\lambda_N^{2d}\sum_{t=1}^{N} Y^2(t-1) - \frac{1}{2}N^{-1}e^{-c/N}\lambda_N^{2d}\sum_{t=1}^{N}(X_{d,\lambda_N}(t))^2 \\
&=: I_{11} - \frac{1}{2}e^{-c/N}N(e^{2c/N}-1)I_{12} + I_{13}.
\end{aligned}$$

Using the continuous mapping theorem and part (i) in Lemma 3.1, we see that

$$(I_{11}, I_{12}) \xrightarrow{\text{d}} \Big(\frac{1}{2}\big(J_c(1)\big)^2, \int_0^1 \big(J_c(s)\big)^2 ds\Big).$$

Employing this result in (5.3), together with $I_{13} \xrightarrow{p} 0$ by Proposition 3.2, we have

$$\begin{aligned}
\min(1, \lambda_N^{-2d})N(\hat{\beta}_N - \beta) &= \frac{I_{11} - \frac{1}{2}N(e^{2c/N}-1)e^{-c/N}I_{12} + I_{13}}{I_{12}} \\
&\xrightarrow{\text{d}} \Big[\int_0^1 (J_c(s))^2\, ds\Big]^{-1}\Big(\frac{1}{2}\Big(J_c(1)\Big)^2 - c\int_0^1 (J_c(s))^2\, ds\Big),
\end{aligned}$$

as required.

□

*Proof of Theorem 4.1.* It suffices to show

(i) the tightness of $\frac{\lambda_i^{d_i}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_i,\lambda_i}(k)$, $i=1,...,m$; and
(ii) the finite dimensional convergence of

$$\mathbf{D}_N\,\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t)=\Big(\frac{\lambda_1^{d_1}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_1,\lambda_1}(k),...,\frac{\lambda_m^{d_m}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_m,\lambda_m}(k)\Big).$$

Let $b_k=e^{-\lambda_1 k}b_{d_1}(k)$, $A_{1,m}=\sum_{j=1}^m \zeta_1(j)\sum_{k=0}^{m-j} b_k$ and $A_{2,m}=\sum_{j=1}^m\sum_{k=0}^\infty b_{k+j}\zeta_1(j)$. Since, for any $0\le t\le 1$,

$$\sum_{k=1}^{[Nt]} X_{d_1,\lambda_1}(k) \;=\; \sum_{k=1}^{[Nt]}\sum_{j=-\infty}^{k} b_{k-j}\zeta_1(j)=A_{1,[Nt]}+A_{2,[Nt]}, \tag{5.4}$$

the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_1,\lambda_1}(k)$ follows from the following proposition, which will be proved in Section 6.

**Proposition 5.1** $\frac{\lambda_1^{d_1}}{\sqrt{N}}A_{1,[Nt]}, 0\le t\le 1$, *is tight and*

$$\mathbb{E}\max_{1\le m\le N}|A_{2,m}| \;=\; o(1)\lambda_1^{-d_1}\sqrt{N}. \tag{5.5}$$

The proof for the tightness of $\frac{\lambda_i^{d_i}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_i,\lambda_i}(k), i=2,...,m$, is similar.

We next prove the finite dimensional convergence of $\mathbf{D}_N\,\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t)$, which easily follows from the following claim: for any fixed $0<t<1$,

$$\frac{\lambda_i^{d_i}}{\sqrt{N}}\sum_{k=1}^{[Nt]} X_{d_i,\lambda_i}(k) \;=\; \frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\zeta_i(k)+o_P(1),\quad i=1,2,...,m. \tag{5.6}$$

due to the the classical result: on $D_{R^m}[0,1]$,

$$\mathbf{S}_N(t):=\Big(\frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\zeta_1(k),\cdots,\frac{1}{\sqrt{N}}\sum_{k=1}^{[Nt]}\zeta_m(k)\Big)\Rightarrow\mathbf{B}(t). \tag{5.7}$$

In fact, by recalling (5.4), we may write (without loss of generality, assume $t=1$ and $i=1$)

$$\begin{aligned}\sum_{k=1}^{N} X_{d_1,\lambda_1}(k) \;&=\; \sum_{k=0}^{N} b_k\sum_{j=1}^{N}\zeta_1(j)+A_{2,N}-\sum_{j=1}^{N}\zeta_1(j)\sum_{k=N-j}^{N} b_k\\ &:=\; \sum_{k=0}^{N} b_k\sum_{j=1}^{N}\zeta_1(j)+A_{2,N}-A_{3,N}.\end{aligned}$$

It is readily seen by using (6.1) of Lemma 6.1 in Section 6 that

$$\begin{aligned}\lambda_1^{d_1}\sum_{k=0}^{N} b_k \;&=\; \frac{\lambda_1^{d_1}}{\Gamma(d_1)}\sum_{k=1}^{N} k^{d_1-1}e^{-\lambda_1 k}+o(1)\\ &=\; 1+o(1)\end{aligned}$$

Similarly, by using (6.5) of Lemma 6.1, we get

$$\mathbb{E}A_{2N}^2+\mathbb{E}A_{3N}^2 \;\le\; 2\sum_{j=0}^{\infty}\Big(\sum_{k=1}^{N} b_{k+j}\Big)^2=o(1)\lambda_1^{-2d_1}N,$$

i.e., $\frac{\lambda_1^{d_1}}{\sqrt{N}}(|A_{2N}|+|A_{3N}|) = o_P(1)$. Combining there facts, we have established (5.6) with $i=1$. The other cases are similar. The proof of Theorem 4.1 is now complete. □

*Proof of Theorem 4.2.* It follows from Theorem 4.1 and Theorem 4.3. □

*Proof of Theorem 4.3.* It only needs to be shown that, for all $1 \le i, j, l \le m$,

$$\Big(\widehat{\mathbf{Y}}(t), \frac{\lambda_i^{d_i}\lambda_j^{d_j}}{N}\sum_{k=1}^{N} X_{d_i,\lambda_i}(k)Y_j(k-1)\Big) \Rightarrow \big(J_C(t), \int_0^1 Q_{js}dB_i(s) + \Delta_{ij}\big), \tag{5.8}$$

jointly on $D_{R^{3m}}[0,1]$. Due to (5.6), we have

$$\mathbf{D}_N\,\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t) \quad = \quad \mathbf{S}_N(t) + o_P(1),$$

where $\mathbf{S}_N(t)$ is defined as in (5.7). On the other hand, $\widehat{\mathbf{Y}}(t) = \mathbf{D}_N\,\mathbf{Y}(t)$ can be presented as a functional of $\mathbf{D}_N\,\mathbf{S}_N^{\mathbf{d},\boldsymbol{\lambda}}(t)$ as in (5.2) (taking $m=0$ and $n=[Ns]$). It is readily seen, by using the same arguments as in the proof of Lemma 3.1, that result (5.8) will follow if we prove

$$\Big(\mathbf{S}_N(t), \frac{\lambda_i^{d_i}\lambda_j^{d_j}}{N}\sum_{k=1}^{N} X_{d_i,\lambda_i}(k)Y_j(k-1)\Big) \Rightarrow \Big(\mathbf{B}(t), \int_0^1 Q_{js}dB_i(t) + \Delta_{ij}\Big), \tag{5.9}$$

jointly on $D_{R^{3m}}[0,1]$.

We only prove (5.9) with $i=2, j=1$ and $m=2$. Due to linearity, extensions to the general $m>2$ case and to joint convergence are straightforward and the details are omitted for brevity.

Let $b_k = e^{-\lambda_1 k} b_{d_1}(k)$ and $c_k = e^{-\lambda_2 k} b_{d_2}(k)$ as in the proof of Theorem 4.1. Recall that

$$X_{d_1,\lambda_1}(k) \quad = \quad \sum_{j=0}^{\infty} b_j u_{k-j},$$

where $u_{k-j} = \zeta_1(k-j),\ b_j \sim \frac{1}{\Gamma(d_1)}\, j^{d_1-1}\, e^{-\lambda_1 j},\ \lambda_1 \equiv \lambda_{1N}$;

$$X_{d_2,\lambda_2}(k) \quad = \quad \sum_{j=0}^{\infty} c_j w_{k-j},$$

where $w_{k-j} = \zeta_2(k-j),\ c_j \sim \frac{1}{\Gamma(d_2)}\, j^{d_2-1}\, e^{-\lambda_2 j},\ \lambda_2 \equiv \lambda_{2N}$; and

$$\begin{aligned} Y_1(k) &= e^{c/N} Y_1(k-1) + X_{d_1,\lambda_1}(k), \quad Y_1(0)=0,\ c\ge 0 \\ &= \sum_{s=1}^{k} e^{(k-s)c/N} X_{d_1,\lambda_1}(k). \end{aligned}$$

As in (3.1)-(3.3), (4.1)-(4.2) and (4.4) of Davidson and Hashimzade (2009), we may write

$$\sum_{t=1}^{N} X_{d_2,\lambda_2}(t)\,Y_1(t-1) \quad =: \quad G_{1N} + G_{2N} + G_{3N},$$

where

$$\begin{aligned} G_{1N} &= \sum_{t=1}^{N-1}\sum_{s=1}^{t} e^{(t-s)c/N} \sum_{m=-\infty}^{t}\ \sum_{i=-\infty}^{\min(s,m)} b_{s-i}c_{t-m}u_i w_{m+1} \\ &= \sum_{m=-\infty}^{N-1} q_{mN} w_{m+1} \end{aligned}$$

with $q_{mN} = \sum_{i=-\infty}^{m} a_{m,i}\, u_i$ and

$$a_{m,i} := a_{m,i}(N) \quad = \quad \sum_{k=\max(1-m,0)}^{N-1-m} c_k \sum_{j=\max(1-i,0)}^{k+m-i} e^{(k+m-i-j)c/N} b_j;$$

$$G_{2N} \quad = \quad \sum_{t=1}^{N-1} \sum_{s=1}^{t} e^{(t-s)c/N} \sum_{k=0}^{\infty} b_k c_{k+t-s+1} u_{s-k} w_{s-k};$$

$$\begin{aligned} G_{3N} \quad &= \quad \sum_{t=1}^{N-1} \sum_{s=1}^{t} e^{(t-s)c/N} \sum_{k=0}^{\infty} \sum_{j=k+t-s+2}^{\infty} b_k c_j u_{s-k} w_{t+1-j} \\ &= \quad \sum_{i=-\infty}^{N-1} h_{i-1,N} u_i, \end{aligned}$$

with $h_{i,N} = \sum_{m=-\infty}^{i} e_{m,i}\, w_m$ and

$$e_{m,i} := e_{m,i}(N) \quad = \quad \sum_{s=\max(1-i,0)}^{N-1-i} b_s \sum_{t=s+i+1-m}^{N-m} e^{(t-s-i-1+m)c/N} c_t.$$

Next let

$$\begin{aligned} \widehat{a}_{m,i} \quad &= \quad e^{(m-i)c/N} \sum_{k=0}^{N-1-m} e^{kc/N} c_k \sum_{j=k+m-i+1}^{N} e^{-jc/N} b_j \\ &\quad + e^{(m-i)c/N} \sum_{k=N-m}^{N-1} e^{kc/N} c_k \sum_{j=0}^{N} e^{-jc/N} b_j, \\ \widetilde{a}_{m,i} \quad &= \quad e^{(m-i)c/N} \sum_{k=0}^{N-1} e^{kc/N} c_k \sum_{j=0}^{N} e^{-jc/N} b_j. \end{aligned}$$

Note that $a_{m,i} = \widetilde{a}_{m,i} - \widehat{a}_{m,i}$. We further have

$$\begin{aligned} G_{1N} \quad &= \quad \sum_{m=-\infty}^{N-1} q_{mN} w_{m+1} \\ &= \quad \sum_{m=1}^{N-1} q_{mN,1} w_{m+1} + \sum_{m=1}^{N-1} q_{mN,2} w_{m+1} + \sum_{m=-\infty}^{0} q_{mN} w_{m+1} \\ &= \quad \sum_{m=1}^{N-1} \widetilde{q}_{mN,1} w_{m+1} - \sum_{m=1}^{N-1} \widehat{q}_{mN,1} w_{m+1} + \sum_{m=1}^{N-1} q_{mN,2} w_{m+1} + \sum_{m=-\infty}^{0} q_{mN} w_{m+1} \\ &=: \quad G_{1N,1} - G_{1N,2} + G_{1N,3} + G_{1N,4}, \end{aligned}$$

where $q_{mN,1} = \sum_{i=1}^{m} a_{m,i}\, u_i$, $q_{mN,2} = q_{mN} - q_{mN,1} = \sum_{i=-\infty}^{0} a_{m,i}\, u_i$, $\widehat{q}_{mN,1} = \sum_{i=1}^{m} \widehat{a}_{m,i}\, u_i$ and

$$\widetilde{q}_{mN,1} = \sum_{i=1}^{m} \widetilde{a}_{m,i}\, u_i = \sum_{k=0}^{N-1} e^{kc/N} c_k \sum_{j=0}^{N} e^{-jc/N} b_j \sum_{i=1}^{m} e^{(m-i)c/N}\, u_i$$

After these preliminaries, result (5.9) with $i = 2$ and $j = 1$ will follow if we prove the following propositions.

**Proposition 5.2** *We have*

$$(5.10)\qquad \Big(\mathbf{S}_N(t),\quad \frac{1}{\sqrt{N}}\sum_{i=1}^{[Nt]} e^{([Nt]-i)c/N}\, u_i\Big)\ \Rightarrow\ \Big(\mathbf{B}(t),\ Q_{1t}\Big),$$

*on $D_{R^3}[0,1]$ in the Skorohod topology, and*

$$(5.11)\qquad \lambda_1^{d_1}\sum_{j=0}^{N} e^{-jc/N} b_j \to 1, \quad \lambda_2^{d_2}\sum_{k=0}^{N-1} e^{kc/N} c_k \to 1.$$

**Proposition 5.3** *We have*

$$(5.12)\qquad N^{-1}\,\lambda_1^{d_1}\lambda_2^{d_2}\big(|G_{1N,2}|+|G_{1N,3}|+|G_{1N,4}|+|G_{3N}|\big) \ =\ o_P(1),$$

**Proposition 5.4** *Suppose that $\mathbb{E}||\zeta(0)||^4<\infty$ and $\lambda_{2N}/\lambda_{1N}\to\eta_{21}$ as $N\to\infty$. We have*

$$(5.13)\qquad N^{-1}\,\lambda_1^{d_1}\lambda_2^{d_2} G_{2N} \ =\ \Delta_{21}+o_P(1).$$

Indeed, by noting that $G_{1N,1}, N\geq 1$, forms a martingale sequence, Proposition 5.2, together an application of Kurtz and Protter (1991) [also see Jacod and Shiryaev (2003)], yield that

$$\Big(\mathbf{S}_N(t),\ \frac{\lambda_1^{d_1}\lambda_2^{d_2}}{N} G_{1N,1}\ \Big)\ \Rightarrow_D\ \Big(\mathbf{B}(t),\ \int_0^1 Q_{1t}dB_2(t)\Big).$$

This result, together with Propositions 5.3 and 5.4, imply the required (5.9) with $i=2$ and $j=1$. The proof of Theorem 4.3 is then complete.

□

The proofs of Propositions 5.2 - 5.4 will be given in next section.

**6. Proofs of Propositions .** Except where mentioned explicitly, the notations are the same as in previous sections. We start with the following lemma, which plays a key role in the proofs of the three propositions.

**Lemma 6.1** *(a) For any $d>0$ and $0<\lambda_n\to\infty$, we have*

$$(6.1)\qquad \Big|\frac{1}{n}\sum_{s=1+[nb]}^{[na]} e^{[\gamma n]/n}\, e^{-\lambda_n s/n}(s/n)^{d-1}-\int_b^a e^{\gamma}\, e^{-\lambda_n u}u^{d-1}du\Big| \ =\ o(1),$$

*uniformly for $0\leq b<a\leq A_0$ for some $A_0<\infty$, as $n\to\infty$.*

*(b) For any $d>0$ and $0<\lambda\equiv\lambda_N\to 0$ satisfying $\lambda N\to\infty$,*

$$(6.2)\qquad \sum_{k=1}^{N} k^{d-1}e^{-\lambda k}\ =\ O(\lambda^{-d}),\quad \sum_{k=N}^{\infty} k^{d-1}e^{-\lambda k}=o(\lambda^{-d})$$

*and unformly for $0\leq s<t\leq 1$,*

$$(6.3)\qquad \sum_{m=0}^{N}\Big(\sum_{k=1+m}^{[Nt]-[Ns]+m} k^{d-1}e^{-\lambda k}\Big)^2\ \leq\ C\,\lambda^{-2d}N(t-s).$$

*(c) For any $d>0$ and $0<\lambda\equiv\lambda_N\to 0$ satisfying $\lambda N\to\infty$, we have*

$$\sum_{m=1}^{N}\Big(\sum_{k=m}^{\infty}k^{2d-2}e^{-2\lambda k}\Big)^{1/2} = o(1)\,\lambda^{-d}\sqrt{N}. \tag{6.4}$$

$$\sum_{m=0}^{\infty}\Big(\sum_{k=1+m}^{N+m}k^{d-1}e^{-\lambda k}\Big)^{2} = o(1)\,\lambda^{-2d}N, \tag{6.5}$$

*Proof.* (6.1) is a well-known result. The proof of result (6.2) is simple. Result (6.3) follows from

$$\begin{aligned}
&\sum_{m=0}^{N}\Big(\sum_{k=1+m}^{[Nt]-[Ns]+m}k^{d-1}e^{-\lambda k}\Big)^{2}\\
\le\;& C\,N^{1+2d}\int_0^1\Big(\int_x^{t-s+x}y^{d-1}e^{-\lambda Ny}dy\Big)^2dx\\
\le\;& C\,\lambda^{-2d}\,N\Big(\int_0^{t-s}+\int_{t-s}^1\Big)\Big(\int_{\lambda Nx}^{\lambda N\{(t-s)+x\}}y^{d-1}e^{-y}dy\Big)^2dx\\
\le\;& C\,\lambda^{-2d}\,N\,(t-s)\Big(\int_0^{\infty}s^{d-1}e^{-s}ds\Big)^2\\
&+C\,\lambda^{-2d}\,N\int_{t-s}^1\Big[e^{-\lambda Nx}(\lambda Nx)^{d-1}\lambda N(t-s)\Big]^2dx\\
\le\;& C_1\,\lambda^{-2d}\,N\,(t-s)+C\,\lambda^{-2d+1}\,N^2(t-s)^2\int_{\lambda N(t-s)}^{\infty}e^{-2x}x^{2(d-1)}dx\\
\le\;& C_1\,\lambda^{-2d}\,N\,(t-s)+C\,\lambda^{-2d}\,N(t-s)\int_{\lambda N(t-s)}^{\infty}e^{-2x}x^{2d-1}dx\\
\le\;& C_2\lambda^{-2d}\,N\,(t-s).
\end{aligned}$$

Similarly, (6.4) follows from

$$\begin{aligned}
&\sum_{m=1}^{N}\Big(\sum_{k=m}^{\infty}k^{2d-2}e^{-2\lambda k}\Big)^{1/2}\le C\sum_{j=1}^{N}\Big(\int_j^{\infty}x^{2d-2}e^{-2\lambda x}dx\Big)^{1/2}\\
\le\;& C\,\lambda^{1/2-d}\sum_{j=1}^{N}\Big(\int_{j\lambda_1}^{\infty}x^{2d-2}e^{-2x}dx\Big)^{1/2}\\
\le\;& C\,\lambda^{1/2-d}N\int_0^1\Big(\int_{\lambda Ny}^{\infty}x^{2d-2}e^{-2x}dx\Big)^{1/2}dy\\
\le\;& C\,\lambda^{-1/2-d}\int_0^{\lambda N}\Big(\int_y^{\infty}x^{2d-2}e^{-2x}dx\Big)^{1/2}dy\\
=\;& o(1)\lambda^{-d}\sqrt{N},
\end{aligned}$$

due to $\lambda N\to\infty$, where we have used the fact:

$$\begin{aligned}
&\int_0^{\infty}\Big(\int_y^{\infty}x^{2d-2}e^{-2x}dx\Big)^{1/2}dy\\
\le\;&\int_0^{\infty}y^{-1/2}e^{-y/2}dy\Big(\int_0^{\infty}x^{2d-1}e^{-x}dx\Big)^{1/2}<\infty.
\end{aligned}$$

We finally prove (6.5). As in the proof of (6.3), we have

$$\begin{aligned}
&\sum_{m=0}^{N}\Big(\sum_{s=1+m}^{N+m} s^{d-1}e^{-\lambda s}\Big)^2 \\
\le\;& C\,N^{1+2d}\int_0^1\Big(\int_x^{1+x} s^{d-1}e^{-\lambda N s}ds\Big)^2dx \\
\le\;& C\,\lambda^{-2d}\,N\Big(\int_0^{1/(\lambda N)^{1/2}}+\int_{1/(\lambda N)^{1/2}}^1\Big)\Big(\int_{\lambda N x}^\infty s^{d-1}e^{-s}ds\Big)^2dx \\
\le\;& C\,\lambda^{-2d}\,N\,(\lambda N)^{-1/2}\Big(\int_0^\infty s^{d-1}e^{-s}ds\Big)^2+C\,\lambda^{-2d}\,N\Big(\int_{(\lambda N)^{1/2}}^\infty s^{d-1}e^{-s}ds\Big)^2 \\
=\;& o(1)\,\lambda^{-2d}\,N, \qquad (6.6)
\end{aligned}$$

as $\lambda N\to\infty$. On the other hand, it is readily seen that

$$\begin{aligned}
\sum_{m=N}^{\infty}\Big(\sum_{s=1+m}^{N+m} s^{d-1}e^{-\lambda s}\Big)^2 \;&\le\; \sum_{m=N}^{\infty} m^{2d}e^{-2\lambda m} \\
&\le\; C\,\lambda^{-2d-1}\int_{\lambda N}^\infty x^{2d}e^{-x}dx = o(1)\,\lambda^{-2d}\,N, \qquad (6.7)
\end{aligned}$$

as $\lambda N\to\infty$. Hence (6.5) follows from (6.6) and (6.7). □

We now turn to the proofs of the propositions. Recall that

$$u_i=\zeta_1(i),\quad w_i=\zeta_2(i),\quad b_j\sim\frac{1}{\Gamma(d_1)}\,j^{d_1-1}\,e^{-\lambda_1 j},\quad c_j\sim\frac{1}{\Gamma(d_2)}\,j^{d_2-1}\,e^{-\lambda_2 j}.$$

*Proof of Proposition 5.1.* It follows from (6.4) that

$$\begin{aligned}
\mathbb{E}\max_{1\le m\le N}|A_{2,m}| \;&\le\; \sum_{j=1}^{N}\mathbb{E}\Big|\sum_{k=0}^\infty b_{k+j}u_j\Big| \le (\mathbb{E}u_0^2)^{1/2}\sum_{j=1}^N\big(\sum_{k=j}^\infty b_k^2\big)^{1/2} \\
&=\; o(1)\times\lambda_1^{-d_1}\sqrt{N},
\end{aligned}$$

i.e., (5.5) is proved. To prove the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}}A_{1,[Nt]}$, we first assume $E|u_0|^{2+\delta}<\infty$ for some $\delta>0$. Since, for any $m_1<m_2$,

$$A_{1,m_2}-A_{1,m_1}=\sum_{j=m_1+1}^{m_2}u_j\sum_{k=0}^{m_2-j}b_k+\sum_{j=1}^{m_1}u_j\sum_{k=m_1+1-j}^{m_2-j}b_k,$$

classical arguments yield [see, for instance, Lemma 1 of Gorodetskii (1977)] that

$$\begin{aligned}
&\mathbb{E}\big|A_{1,[Nt_2]}-A_{1,[Nt_1]}\big|^{2+\delta} \\
\le\;& CE|u_0|^{2+\delta}\Big(\sum_{j=[Nt_1]+1}^{[Nt_2]}\big[\sum_{k=0}^{[Nt_2]-j}b_k\big]^2+\sum_{j=0}^{[Nt_1]-1}\big[\sum_{k=j+1}^{[Nt_2]-[Nt_1]+j}b_k\big]^2\Big)^{(2+\delta)/2} \\
\le\;& C\,\big(\sqrt{N}/\lambda_1^{d_1}\big)^{2+\delta}\,(t_2-t_1)^{1+\delta/2},
\end{aligned}$$

for any $0 \le t_1 < t_2 \le 1$, due to (6.2) and (6.3). This yields the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,[Nt]}, 0 \le t \le 1$, by Theorem 15.6 of Billingsley (1968).

We next prove the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,[Nt]}$ without the restriction: $E|u_0|^{2+\delta} < \infty$ for some $\delta > 0$. In fact, by Major (1976), we may redefine $\{u_k, k \ge 1\}$ on a richer probability space together with a sequence of independent normal random variables $\{Y_k, k \ge 1\}$ with $\mathbb{E}Y_1 = 0$ and $\mathbb{E}Y_1^2 = \sigma_1^2$ such that for all $\epsilon > 0$,

$$P\Big(\max_{1 \le k \le N} \big|S_k - Z_k\big| \ge \epsilon\sqrt{N}\Big) \to 0, \tag{6.8}$$

as $N \to \infty$, where $S_k = \sum_{j=1}^k u_j$ and $Z_k = \sum_{j=1}^k Y_j$. Result (6.8), together with (6.2), implies the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,[Nt]}$. Indeed, by letting $Z_{N,m} = \frac{\lambda_1^{d_1}}{\sqrt{N}} \sum_{j=1}^m Y_j \sum_{k=0}^{m-j} b_k$, we have

$$\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,m} - Z_{N,m} = \frac{\lambda_1^{d_1}}{\sqrt{N}} \sum_{k=1}^m b_{m-k}\big(S_k - Z_k\big)$$

for any $1 \le m \le N$. Since $Z_{N,[Nt]}, 0 \le t \le 1$, is tight as proved above, the tightness of $\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,[Nt]}$ follows from

$$\max_{1 \le m \le N} \Big|\frac{\lambda_1^{d_1}}{\sqrt{N}} A_{1,m} - Z_{N,m}\Big| \quad \le \quad C\,\frac{1}{\sqrt{N}} \max_{1 \le m \le N} |S_k - Z_k|\, \lambda_1^{d_1} \sum_{k=1}^N b_k = o_P(1),$$

due to (6.8) and (6.2). The proof of Proposition 5.1 is now complete. □

*Proof of Proposition 5.2.* The proof of (5.10) is similar to that of Lemma 3.1 but simpler. The proof of (5.11) is similar to (6.9) below and the details are omitted. □

*Proof of Proposition 5.3.* We only prove $N^{-1}\lambda_1^{d_1}\lambda_2^{d_2}|G_{2N}| = o_P(1)$ . The other results are similar but simpler. By using the independence of $(u_k, w_k)$, we have

$$\begin{aligned}
\mathbb{E}G_{3N}^2 &= \sum_{i=-\infty}^{N-1} \mathbb{E}h_{i-1,N}^2 \mathbb{E}u_1^2 \\
&= \sum_{i=-\infty}^{N-1} \sum_{m=-\infty}^{i-1} e_{m,i-1}^2 \mathbb{E}u_1^2 \mathbb{E}w_1^2 \\
&\le C \sum_{i=-\infty}^{N-1} \sum_{m=-\infty}^{i-1} \Big( \sum_{s=\max(1-i,0)}^{N-1-i} b_s \sum_{t=s+i+1-m}^{N-m} c_t \Big)^2 \\
&\le C \sum_{i=1}^{N-1} \sum_{m=-\infty}^{i-1} \Big( \sum_{s=0}^{N-1-i} b_s \sum_{t=s+i+1-m}^{N-m} c_t \Big)^2 + C \sum_{i=0}^{\infty} \sum_{m=i+1}^{\infty} \Big( \sum_{s=1+i}^{N+i} b_s \sum_{t=s-i+m}^{N+m} c_t \Big)^2 \\
&\le C \Big( \sum_{s=0}^{N} b_s \Big)^2 \sum_{i=1}^{N} \sum_{m=1}^{\infty} \Big( \sum_{t=1+m}^{N+m-i} c_t \Big)^2 \quad \text{(by using transformation } i - m \to m) \\
&\quad + C \sum_{i=0}^{\infty} \sum_{m=i+1}^{\infty} \Big( \sum_{s=1+i}^{N+i} b_s \sum_{t=1+m}^{N+m} c_t \Big)^2 \\
&\le C \Big[ N \Big( \sum_{s=0}^{N} b_s \Big)^2 + \sum_{i=0}^{\infty} \Big( \sum_{s=1+i}^{N+i} b_s \Big)^2 \Big] \sum_{m=0}^{\infty} \Big( \sum_{t=1+m}^{N+m} c_t \Big)^2.
\end{aligned}$$

Now, it follows from (6.2) and (6.5) of Lemma 6.1 that

$$\mathbb{E}G_{3N}^2 \quad = \quad o(1) \times N^2 \lambda_1^{-2d_1} \lambda_2^{-2d_2},$$

i.e., $N^{-1}\lambda_1^{d_1}\lambda_2^{d_2}|G_{3N}| = o_P(1)$ as required. □

*Proof of Proposition 5.4.* Write

$$A_N \quad = \quad \sum_{t=1}^{N-1} \sum_{s=1}^{t} e^{(t-s)c/N} \sum_{k=0}^{N} b_k c_{k+t-s+1}.$$

By recalling the definition of $b_k$ and $c_k$, it follows from (6.1) and $\lambda_1 N \to \infty$ and $\lambda_2 N \to \infty$ that

$$\begin{aligned}
A_N &= \sum_{s=1}^{N-1} \sum_{k=0}^{N} b_k \sum_{t=0}^{N-1-s} e^{tc/N} c_{k+t+1} \\
&\sim \frac{N^{1+d_1+d_2}}{\Gamma(d_1)\Gamma(d_2)} \int_0^1 \int_0^1 \int_0^{1-s} x^{d_1-1} e^{-\lambda_1 N x} e^{yc} (y+x)^{d_2-1} e^{-\lambda_2 N (y+x)} dydxds \\
&\sim \frac{N\lambda_1^{-d_1}\lambda_2^{-d_2}}{\Gamma(d_1)\Gamma(d_2)} \int_0^1 \int_0^{\lambda_1 N} \int_{\lambda_2 x/\lambda_1}^{\lambda_2 N(1-s+x)} x^{d_1-1} e^{-x} e^{yc/\lambda_2 N} y^{d_2-1} e^{-y} dydxds \\
&\sim \frac{N\lambda_1^{-d_1}\lambda_2^{-d_2}}{\Gamma(d_1)\Gamma(d_2)} \int_0^\infty \int_{\lambda_2 x/\lambda_1}^{\infty} x^{d_1-1} e^{-x} y^{d_2-1} e^{-y} dydx
\end{aligned}$$

$$(6.9) \qquad \sim N\lambda_1^{-d_1}\lambda_2^{-d_2} \begin{cases} 1, & \text{if } \lambda_2/\lambda_1 \to 0, \\ \frac{1}{\Gamma(d_1)\Gamma(d_2)} \int_0^\infty \int_{\eta x}^\infty x^{d_1-1} e^{-x} y^{d_2-1} e^{-y} dydx, & \\ \qquad\qquad \text{if } \lambda_2/\lambda_1 \to 0 < \eta_{21} < \infty & \\ o(1), & \text{if } \lambda_2/\lambda_1 \to \infty. \end{cases}$$

This, together with the fact that

$$\begin{aligned}
|\mathbb{E}G_{2N} - A_N \mathbb{E}u_1 w_1| &\leq \mathbb{E}(u_1 w_1) \sum_{t=1}^{N-1} \sum_{s=1}^{t} \sum_{k=N+1}^{\infty} b_k c_{k+t-s+1} \\
&\leq CN \sum_{k=N+1}^{\infty} k^{d_1-1} e^{-\lambda_1 k} \sum_{k=N+1}^{\infty} k^{d_2-1} e^{-\lambda_2 k} = o(1) \times N\lambda_1^{-d_1}\lambda_2^{-d_2},
\end{aligned}$$

due to (6.2), yields

$$N^{-1}\lambda_1^{d_1}\lambda_2^{d_2}\mathbb{E}G_{2N} = A_N \mathbb{E}\big[\zeta_1(0)\zeta_2(0)\big] = \Delta_{21} + o(1)$$

Result (5.13) will follow if we prove

$$(6.10) \qquad G_{2N} - \mathbb{E}G_{2n} \quad = \quad o_P\big[N\lambda_1^{-d_1}\lambda_2^{-d_2}\big].$$

In fact, by noting

$$\begin{aligned}
G_{2N} - \mathbb{E}G_{2n} &= \sum_{t=1}^{N-1} \sum_{s=1}^{t} e^{(t-s)c/N} \sum_{k=-\infty}^{s} b_{s-k} c_{t-k+1} \eta_k \\
&= \sum_{k=-\infty}^{N-1} \eta_k \sum_{t=\max\{1,k\}}^{N-1} \sum_{s=\max\{1,k\}}^{t} e^{(t-s)c/N} b_{s-k} c_{t-k+1},
\end{aligned}$$

where $\eta_k = u_k w_k - \mathbb{E}(u_k w_k)$, we have

$$\begin{aligned}\mathbb{E}\big(G_{2N} - \mathbb{E}G_{2n}\big)^2 &\leq \mathbb{E}\eta_1^2 \sum_{k=1}^{N-1}\Big(\sum_{t=1}^{N} c_{t-k+1}\Big)^2\big(\sum_{s=1}^{t} b_{s-k}\big)^2 + \sum_{k=0}^{\infty}\Big(\sum_{t=1}^{N-1} c_{t+k+1}\Big)^2\Big(\sum_{s=1}^{t} b_{s+k}\Big)^2 \\ &\leq CN\,\lambda_1^{-2d_1}\lambda_2^{-2d_2},\end{aligned}$$

due to (6.2) - (6.5) and $\mathbb{E}\eta_0^2 \leq 4(\mathbb{E}u_0^4)^{1/2}(\mathbb{E}w_0^4)^{1/2} < \infty$. This yields (6.10). The proof fo Proposition 5.4 is complete. $\square$

## REFERENCES


[1] Abramowitz, M., Stegun, I.: Handbook of mathematical functions, ninth edition, Dover, New York (1965).

[2] Ahtola, J. & Tiao, G. C. (1984). Parameter inference for a nearly nonstationary first order autoregressive model. *Biometrika* **71**, 263–72.

[3] Taqqu, M. S., Bai, S. (2013). Multivariate limits of multilinear polynomial-form processes with long memory. *Statistics and Probability Letters* **83**, 2473–2485

[4] Beran, J. (1994). *Statistics for long-memory processes*, Chapman and Hall/CRC.

[5] Beran, J., Yuanhua, F., Ghosh, S. and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer.

[6] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

[7] Bolin, D. and Lindgren, F. (2011b) Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. Ann. Appl. Statist., 5, 523-550

[8] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. New York: Springer.

[9] Buchmann, B. and Chan, N. H. (2007). Asymptotic theory of least squares estimators for nearly unstable processes under strong dependence. *Annals of Statistics* **35** 2001-2017. MR2363961

[10] Chan, N.H. and C.Z. Wei. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* **15** 1050–1063.

[11] Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein–Uhlenbeck processes. *Electron. J. Probab* **8** 1–14.

[12] Chung, C. F.(2002). Sample Means, Sample Autocovariances, and Linear Regression of Stationary Multivariate Long Memory Processes *Econometric Theory* **18**, No. 1, 51–78

[13] Davidson, J., DE Jong, R. M.(2002). The Functional Central Limit Theorem and Weak Convergence to Stochastic Integrals II: Fractionally Integrated Processes. *Econometric Theory* **16**, No. 5, 643–666

[14] Cartea, Á., del-Castillo-Negrete, D. (2007). Fluid limit of the continuous-time random walk with general Lévy jump distribution functions. *Physical Review E* **76**, 041105 .

[15] Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *JASA* **74** 427–431.

[16] Embrechts, P., Maejima, M., (2002). *Selfsimilar Processes*. Princeton University Press.

[17] Evans, G. B. A., Savin, N. E. (1981). Testing for unit roots: 1. *Econometrica* **49**, 753–79.

[18] Evans, G. B. A., Savin, N. E. (1984). Testing for unit roots: 2. *Econometrica* **52**, 1241–69.

[19] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory* **16** 3–22.

[20] Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003). Rescaled variance and related tests for long memory in volatility and levels. *J. Econometrics* **112** 265–294.

[21] Giraitis, L., Koul. H.L. and Surgailis, D. (2012). *Large Sample Inference for Long Memory Processes.* London: Imperial College Press.

[22] Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Tables of Integrals and Products.* 6th edition. New York: Academic Press.

[23] Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Series Anal.* **1** 15–29.

[24] Granger, C.W.J. and Ding, Z (1996). Varieties of long memory models. *J. Econometrics* **73** 61–77.

[25] Hosking, J.R.M. (1981). Fractional differencing. *Biometrika* **68** 165–176.

[26] Kasahara, Y. and Maejima, M. (1988). Weighted sums of i.i.d. random variables attracted to integrals of stable processes. *Probab. Theory Relat. Fields* **78** 75–96.

[27] Kokoszka, P.S. and Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stochastic Process. Appl.* **60** 19–47.

[28] Lavancier, F., Leipus, R., Philippe, A. and Surgailis, D. (2013). Detection of non-constant long memory parameter. *Econometric Theory* **29** 1009–1056.

[29] Lo, A. (1991). Long-term memory in stock market prices. *Econometrica* **59** 1279–1313.

[30] Meerschaert, M.M. and Sabzikar, F. (2014). Stochastic integration with respect to tempered fractional Brownian motion. *Stochastic Process. Appl.* **124** 2363–2387.

[31] Meerschaert, M.M. and Sabzikar, F. (2013). Tempered fractional Brownian motion. *Statist. Probab. Lett.* **83** 2269–2275.

[32] Meerschaert, M.M. and Sabzikar, F. (2016). Tempered fractional stable motion. *J. Theoret. Probab.* **29** 681–706.

[33] Meerschaert, M.M., Sabzikar, F., Phanikumar, M.S. and Zeleke, A. (2014). Tempered fractional time series model for turbulence in geophysical flows. *J. Stat. Mech. Theory Exp.* **2014** P09023.

[34] Meerschaert, M.M. and Sikorskii, A. (2012). *Stochastic Models for Fractional Calculus.* Berlin/Boston: De Gruyter.

[35] Peng, J. and Wang, Q. (2018). Weak convergence to stochastic integrals under primitive conditions in nonlinear econometric models.

[36] Pipiras, V. and Taqqu, M. (2000). Integration questions related to fractional Brownian motion, *Probab. Theory Related Fields* **118**, 251–291.

[37] Phillips, P. C. B. (1987a). Time series regression with a unit root. *Econometrica* **55** 277–301.

[38] Phillips, P. C. B. (1987b). Towards a unified asymptotic theory of autoregression. *Biometrika* **74**, 535–48.

[39] Phillips, P. C. B. (1987c). Asymptotic Expansions in Nonstationary Vector Autoregressions. *Econometric Theory* **3** No. 1, 45–68.

[40] Phillips, P. C. B. (1988). Regression theory for near-integrated time series. *Econometrica* **56** No. 5, 1021-1044.

[41] Phillips, P. C. B. (2010). Long memory and long run variation. *Journal of Econometrics* **151**, 150-158.

[42] Phillips, P. C. B. (1988). Weak convergence of sample covariance matrices to stochastic integrals via martingale approximations. *Econometric Theory* **4**, 528533.

[43] Phillips, P. C. B. and Kim, S. C. (2007): Long-Run Covariance Matrices for Fractionally Integrated Processes. *Econometric Theory* **23**, No. 6, 1233–1247.

[44] Sabzikar, F. Meerschaert, M. M. and Chen, J. (2015). Tempered Fractional Calculus.*Journal of Computational Physics*, **293**, 14–28

[45] Sabzikar, F. and Surgailis, D. (2017). Invariance principles for tempered fractionally integrated processes. *Stochastic Processes and their Applications*, DOI10.1016/j.spa.2017.11.004.

[46] Sabzikar, F. and Surgailis, D. (2018). Tempered fractional Brownian and stable motions of second kind. *Statistics & Probability Letters*, **132**, 17–27.

[47] Samorodnitsky, G. and Taqqu, M.S. (1996). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* Boca Raton etc: Chapman and Hall.

[48] Sowell, F. (1990). The fractional unit root distribution. *Econometrica* **58** 495–505.

[49] Wang, Q. (2015). *Limit theorems for nonlinear cointegration regression.* World Scientific.

[50] Zeng, C., Yang, Q., Chen, Y. (2016): Bifurcation dynamics of the tempered fractional Langevin equation, *Chaos* **26**, 084310 ); doi: 10.1063/1.4959533.

[51] Young, L. C. (1936). An inequality of the H¨older type, connected with Stieltjes integration. *Acta Math* **67** 251-282

# APPENDIX

**A. Stochastic integration with respect to TFBM II .** In this section, we define the stochastic integral of a non-random function $f$ with respect to TFBM II by applying the connection between tempered fractional calculus and TFBM II. Recall from [30] that the (positive and negative) tempered fractional integrals (TFI) and tempered fractional derivatives (TFD) of a function $f:\mathbb{R}\to\mathbb{R}$ are defined by

$$\mathbb{I}_{\pm}^{\kappa,\lambda}f(y):=\frac{1}{\Gamma(\kappa)}\int f(s)(y-s)_{\pm}^{\kappa-1}\mathrm{e}^{-\lambda(y-s)_{\pm}}\mathrm{d}s,\quad \kappa>0 \tag{A.1}$$

and

$$\mathbb{D}_{\pm}^{\kappa,\lambda}f(y):=\lambda^{\kappa}f(y)+\frac{\kappa}{\Gamma(1-\kappa)}\int(f(y)-f(s))(y-s)_{\pm}^{-\kappa-1}\,\mathrm{e}^{-\lambda(y-s)_{\pm}}\mathrm{d}s,\quad 0<\kappa<1, \tag{A.2}$$

respectively. The TFI in (A.1) exists a.e. in $\mathbb{R}$ for each $f\in L^p(\mathbb{R})$ and defines a bounded linear operator in $L^p(\mathbb{R}), p\geq 1$ ([30], Lemma 2.2). The TFD in (A.2) exists for any absolutely continuous function $f\in L^1(\mathbb{R})$ such that $f'\in L^1(\mathbb{R})$; moreover, it can be extended to the fractional Sobolev space

$$W^{\kappa,2}(\mathbb{R}):=\left\{f\in L^2(\mathbb{R}):\int(\lambda^2+\omega^2)^{\kappa}|\hat{f}(\omega)|^2\,\mathrm{d}\omega<\infty\right\}, \tag{A.3}$$

where $\hat{f}$ denotes the Fourier transform of $f$. See ([30], Theorem 2.9 and Definition 2.11).

The following proposition shows that TFBM II can be written as a stochastic integral of TFI/TFD of the indicator function of the interval $[0,t]$. We refer the reader to see [46] for the details. For $t<0$, let $\mathbf{1}_{[0,t]}(y):=-\mathbf{1}_{[-t,0]}(y), y\in\mathbb{R}$.

**Proposition A.1** Let $d>-\frac{1}{2},\lambda>0$, and $t\in\mathbb{R}$. Then

$$B_{d,\lambda}^{II}(t)=\Gamma(d+1)\begin{cases}\int\mathbb{I}_{-}^{d,\lambda}\mathbf{1}_{[0,t]}(y)B(\mathrm{d}y), & d>0,\\ \int\mathbb{D}_{-}^{-d,\lambda}\mathbf{1}_{[0,t]}(y)B(\mathrm{d}y), & -\frac{1}{2}<d<0.\end{cases} \tag{A.4}$$

Now we discuss a general construction for stochastic integrals of non-random functions with respect to TFBM II. For a standard Brownian motion $\{B(t)\}_{t\in\mathbb{R}}$ on $(\Omega,\mathcal{F},P)$, the stochastic integral $\mathcal{I}(f):=\int f(x)B(dx)$ is defined for any $f\in L^2(\mathbb{R})$, and the mapping $f\mapsto\mathcal{I}(f)$ defines an isometry from $L^2(\mathbb{R})$ into $L^2(\Omega)$, called the *Itô isometry*:

$$\langle\mathcal{I}(f),\mathcal{I}(g)\rangle_{L^2(\Omega)}=\mathrm{Cov}[\mathcal{I}(f),\mathcal{I}(g)]=\int f(x)g(x)\,dx=\langle f,g\rangle_{L^2(\mathbb{R})}. \tag{A.5}$$

Define $\mathcal{E}$ as the space of elementary functions

$$f(u)=\sum_{i=1}^{n}a_i\mathbf{1}_{[t_i,t_{i+1})}(u), \tag{A.6}$$

where $a_i,t_i$ are real numbers such that $t_i<t_j$ for $i<j$. It is natural to define the stochastic integral

$$\mathcal{I}^{d,\lambda}(f)=\int_{\mathbb{R}}f(x)B_{d,\lambda}^{II}(dx)=\sum_{i=1}^{n}a_i\Big[B_{d,\lambda}^{II}(t_{i+1})-B_{d,\lambda}^{II}(t_i)\Big]. \tag{A.7}$$

Now, assume $d > 0$. It follows immediately from Proposition A.1 that for $f \in \mathcal{E}$, the stochastic integral

$$\mathcal{I}^{d,\lambda}(f) = \int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) = \int_{\mathbb{R}} \left(\mathbb{I}^{d,\lambda}_{-} f\right)(x)\ B(dx)$$

is a Gaussian random variable with mean zero, such that for any $f, g \in \mathcal{E}$ we have

$$\begin{aligned} \langle \mathcal{I}^{d,\lambda}(f), \mathcal{I}^{d,\lambda}(g)\rangle_{L^2(\Omega)} &= \mathbb{E}\Big( \int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) \int_{\mathbb{R}} g(x) B^{II}_{d,\lambda}(dx)\Big) \\ &= \int_{\mathbb{R}} \left(\mathbb{I}^{d,\lambda}_{-} f\right)(x) \left(\mathbb{I}^{d,\lambda}_{-} g\right)(x)\ dx, \end{aligned} \tag{A.8}$$

in view of (A.4), when $d > 0$, and the Itô isometry (A.5).

Based on (A.8), we define the following class of functions:

**Definition A.2**

$$\mathcal{A}_1 := \left\{ f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} \left| \left(\mathbb{I}^{d,\lambda}_{-} f\right)(x) \right|^2 dx < \infty \right\}, \tag{A.9}$$

*for $d > 0$ and $\lambda > 0$.*

**Theorem A.3** *Given $d > 0$ and $\lambda > 0$, the class of functions $\mathcal{A}_1$, defined by* (A.9)*, is a linear space with the inner product*

$$\langle f, g \rangle_{\mathcal{A}_1} = \int_{\mathbb{R}} \left(\mathbb{I}^{d,\lambda}_{-} f\right)(x) \left(\mathbb{I}^{d,\lambda}_{-} g\right)(x)\ dx \tag{A.10}$$

*The set of elementary functions $\mathcal{E}$ is dense in the space $\mathcal{A}_1$.*

We omit the proof of Theorem A.3 since it is similar to [30, Theorem 3.5].

We now define the stochastic integral with respect to TFBMII for any function in $\mathcal{A}_1$ in the case where $d > 0$.

**Definition A.4** *For any $d > 0$ and $\lambda > 0$, we define*

$$\int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) := \int_{\mathbb{R}} \left(\mathbb{I}^{d,\lambda}_{-} f\right)(x)\ B(dx) \tag{A.11}$$

*for any $f \in \mathcal{A}_1$.*

Next we investigate stochastic integrals with respect to TFBMII in the case $-\frac{1}{2} < d < 0$. It follows from (A.4) that the stochastic integral (A.7) can be written in the form

$$\mathcal{I}^{d,\lambda}(f) = \int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) = \int_{\mathbb{R}} \mathbb{D}^{-d,\lambda}_{-} f(x)\ B(dx)$$

for any $f \in \mathcal{E}$. Then $\mathcal{I}^{d,\lambda}(f)$ is a Gaussian random variable with mean zero, such that

$$\begin{aligned} \langle \mathcal{I}^{d,\lambda}(f), \mathcal{I}^{d,\lambda}(g)\rangle_{L^2(\Omega)} &= \mathbb{E}\Big( \int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) \int_{\mathbb{R}} g(x) B^{II}_{d,\lambda}(dx)\Big) \\ &= \int_{\mathbb{R}} \left(\mathbb{D}^{-d,\lambda}_{-} f\right)(x) \left(\mathbb{D}^{-d,\lambda}_{-} g\right)(x)\ dx \end{aligned} \tag{A.12}$$

for any $f, g \in \mathcal{E}$, using (A.7) and the Itô isometry (A.5). Equation (A.12) suggests the following space of integrands for TFBM II in the case $-\frac{1}{2} < d < 0$.

**Definition A.5**

$$\mathcal{A}_2 := \left\{ f : \varphi_f = \mathbb{D}_-^{-d,\lambda} f \text{ for some } \varphi_f \in L^2(\mathbb{R}) \right\}. \tag{A.13}$$

*for any* $-\frac{1}{2} < d < 0$.

**Theorem A.6** *Given* $-\frac{1}{2} < d < 0$ *and* $\lambda > 0$, *the class of functions* $\mathcal{A}_2$, *defined by* (A.13), *is a linear space with the inner product*

$$\langle f, g \rangle_{\mathcal{A}_2} = \int_{\mathbb{R}} \left(\mathbb{D}_-^{-d,\lambda} f\right)(x) \left(\mathbb{D}_-^{-d,\lambda} g\right)(x)\, dx \tag{A.14}$$

*The set of elementary functions* $\mathcal{E}$ *is dense in the space* $\mathcal{A}_2$.

We omit the proof of Theorem A.6 since it is similar to [30, Theorem 3.10].

We now define the stochastic integral with respect to TFBM II for any function in $\mathcal{A}_2$ in the case where $-\frac{1}{2} < d < 0$.

**Definition A.7** *For any* $-\frac{1}{2} < d < 0$ *and* $\lambda > 0$, *we define*

$$\int_{\mathbb{R}} f(x) B^{II}_{d,\lambda}(dx) := \int_{\mathbb{R}} \left(\mathbb{D}_-^{-d,\lambda} f\right)(x)\ B(dx) \tag{A.15}$$

*for any* $f \in \mathcal{A}_2$.

**B. Tempered fractional linear processes.** This section outlines the univariate ARTFIMA class of processes, introduces the vector autoregressive tempered fractionally moving average (VARTFIMA) class, and discusses some of its properties.

The univariate ARTFIMA $(p, d, \lambda, q)$ was introduced and discussed in [44] based on tempered fractional difference operator. Here we recall some definitions and primary properties of ARTFIMA$(p, d, \lambda, q)$ class in the univariate case. A tempered fractional difference operator is defined by

$$\Delta^{d,\lambda} f(x) = (I - e^{-\lambda} B)^d f(x) = \sum_{j=0}^{\infty} \omega_{d,\lambda}(j) f(x - j) \tag{B.1}$$

where $d > 0$, $\lambda > 0$, and

$$\omega_{d,\lambda}(j) := (-1)^j \binom{d}{j} e^{-\lambda j} \quad \text{where} \quad \binom{d}{j} = \frac{\Gamma(1+d)}{j!\Gamma(1+d-j)} \tag{B.2}$$

using the gamma function $\Gamma(d) = \int_0^\infty e^{-x} x^{d-1}\, dx$. Using the recurrence property $\Gamma(d+1) = d\Gamma(d)$, we can extend (B.1) to non-integer values of $d < 0$. By a common abuse of notation, we call this a tempered fractional integral.

If $\lambda = 0$, then equation (B.1) reduces to the usual fractional difference operator. See [34, 44] for more details.

**Definition B.1** The discrete time stochastic process $\{X_t\}_{t \in \mathbb{Z}}$ is called an *autoregressive tempered fractional integrated moving average* time series, denoted by ARTFIMA$(p, \lambda, d, q)$, if $\{X_t\}$ is a stationary solution with zero mean of the tempered fractional difference equations

$$\Phi(B) \Delta^{d,\lambda} X_t = \Theta(B) \zeta_t, \tag{B.3}$$

where $Z_t$ is a white noise sequence (i.i.d. with $\mathbb{E}[\zeta_t] = 0$ and $\mathbb{E}[\zeta_t^2] = \sigma^2$), $d \notin \mathbb{Z}$, $\lambda > 0$, and $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p$, and $\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q$ are polynomials of degrees $p, q \geq 0$ with no common zeros.

**Remark B.2** Assuming polynomials $\Phi(\cdot)$ and $\Theta(\cdot)$ have no common zeros and

$$\big|\Phi(z)\big| > 0 \quad \text{and} \quad \big|\Theta(z)\big| > 0 \tag{B.4}$$

for $|z| \leq 1$, it can be shown that the ARTFIMA$(p, d, \lambda, q)$ process is causal and invertible.

**Remark B.3** Another version of tempered fractionally integrated process was defined in [45] as follows: The discrete time stochastic process $\{X_t^*\}_{t\in\mathbb{Z}}$ is called ARTFIMA$(p, d, \lambda, q)$ process with innovation process $Z(t)$ if

$$X_t^* = \sum_{k=0}^{\infty} e^{-\lambda k} a_{-d}(k)\zeta_{t-k}, \qquad t \in \mathbb{Z} \tag{B.5}$$

where the coefficients $a_d(k)$ are the coefficients of ARFIMA$(p, d, q)$. That is

$$a_d(k) = \sum_{s=0}^{k} \omega_d(k)\psi(k - s), \tag{B.6}$$

where $\omega_d(k) = \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(d)}$, and $\psi(j)$ are the coefficients of the power series $\sum_{j=0}^{\infty} \psi(j)z^j = \Theta(z)/\Phi(z), |z| \leq 1$.
(ii) When $p = q = 0$, $X_t$ and $X_t^*$ are the same time series. However, in general, they are different stochastic processes. For instance, $X_t$ has the spectral density $f_X(\nu) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-i\nu})\right|^2}{\left|\Phi(e^{-i\nu})\right|^2} \left|1 - e^{-(\lambda+i\nu)}\right|^{-2d}$ for $-\pi \leq \nu \leq \pi$, while $X_t^*$ has the spectral density $f_X(\nu) = \frac{\sigma^2}{2\pi} \frac{\left|\Theta(e^{-(\lambda+i\nu)})\right|^2}{\left|\Phi(e^{-(\lambda+i\nu)})\right|^2} \left|1 - e^{-(\lambda+i\nu)}\right|^{-2d}$ for the same range of $\nu$.

We now proceed to define the vector ARTFIMA model. First, let $\mathbf{X}(t)$ be a real-valued covariance stationary $m$-vector time series generated by the following model:

$$\begin{pmatrix} (1 - e^{-\lambda_1}B)^{d_1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & (1 - e^{-\lambda_m}B)^{d_m} \end{pmatrix} \begin{pmatrix} X_{1t} - \mathbb{E}X_{1t} \\ \vdots \\ X_{mt} - \mathbb{E}X_{mt} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{mt} \end{pmatrix}, \tag{B.7}$$

where $d_1, \ldots d_m$, $\lambda_1, \ldots, \lambda_m$ are the memory and tempering parameters respectively, $B$ is the lag operator, and $\mathbf{u}_t = (u_{1t}, \ldots, u_{mt})'$ is a covariance stationary process. By assuming $\mathbf{u}_t$ is a vector autoregressive integrated moving average (VARIMA) process, we can define a vector autoregressive tempered fractionally integrated moving average (VARTFIMA) process as follows. Suppose $\mathbf{u}_t = (\mathbf{\Phi}(B))^{-1}\mathbf{\Theta}(B)\boldsymbol{\zeta}(t)$, where $\mathbf{\Phi}(B) = \mathbf{\Phi}_0 - \sum_{i=1}^{p} \mathbf{\Phi}_i B^i$ and $\mathbf{\Theta}(B) = \mathbf{\Theta}_0 + \sum_{i=1}^{q} \mathbf{\Theta}_i B^i$ are $(m \times m)$ matrix polynomials in $B$. A VARTFIMA model is defined by

$$\mathbf{\Phi}(B)\mathbf{\Delta}^{d,\lambda}(B)(\mathbf{X}(t) - \boldsymbol{\mu}) = \mathbf{\Theta}(B)\boldsymbol{\zeta}(t),$$

where $\boldsymbol{\mu} = (\mathbb{E}X_{1t}, \ldots, \mathbb{E}X_{mt})' = (\mu_1, \ldots, \mu_m)'$ is the $m \times 1$ mean vector, $\boldsymbol{\zeta}(t)$ is m-dimensional vector with $\mathbb{E}(\boldsymbol{\zeta}(t)) = \mathbf{0}$ and covariance matrix $\mathbf{\Omega}$. The operator $\mathbf{\Delta}^{d,\lambda}(B)$ is the $m \times m$ diagonal matrix given by (B.7).

The following remark gives the autocovariance function of $\mathbf{X}(t)$ and its asymptotic form for large lags when $p = q = 0$.

**Remark B.4** (i) If $d_a \in \mathbb{R} \setminus \mathbb{N}_-$ and $\lambda_a > 0$ for all $a = 1, \dots, m$ and the spectral density matrix $f_{uu}(\omega)$ of $u_t$ is continuously differentiable, then

$$[\Gamma_{xx}]_{ab} = \frac{2 f_{u_a u_b}(0) e^{-\lambda_b k} \Gamma(k + d_b) \; {}_2F_1(d_a, k + d_b; k + 1; \mathrm{e}^{-(\lambda_a + \lambda_b)})}{\Gamma(k + 1)\Gamma(d_b)}. \tag{B.8}$$

(ii) As $k \to \infty$, we have

$$[\Gamma_{xx}]_{ab} \sim K_{ab} \; e^{-\lambda_b k} k^{d_b - 1}, \tag{B.9}$$

where $K_{ab} = \frac{2 f_{u_a u_b}(0) \left(1 - e^{-(\lambda_a + \lambda_b)}\right)^{-d_a}}{\Gamma(d_b)}$.
(iii) Assuming $\lambda_a = \lambda_b = 0$ in (B.8), we have the specialization

$$[\Gamma_{xx}]_{ab} \sim \frac{2 f_{u_a u_b}(0) \Gamma(1 - d_a - d_b) \sin \pi d_b}{k^{1 - d_a - d_b}}, \quad k \to \infty,$$

which is the asymptotic behavior of the elements of the autocovariance matrix in the untempered case, see [41, Section 2.1] or [43, Theorem 1].