Probability in the Engineering and Informational Sciences, **36**, 2022, 144–168.
doi:10.1017/S0269964820000388

BROWNIAN MOTION MINUS THE INDEPENDENT INCREMENTS: REPRESENTATION AND QUEUING APPLICATION

Kerry Fendick
Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA
E-mail: kerry.fendick@jhuapl.edu

This paper relaxes assumptions defining multivariate Brownian motion (BM) to construct processes with dependent increments as tractable models for problems in engineering and management science. We show that any Gaussian Markov process starting at zero and possessing stationary increments and a symmetric smooth kernel has a parametric kernel of a particular form, and we derive the unique unbiased, jointly sufficient, maximum-likelihood estimators of those parameters. As an application, we model a single-server queue driven by such a process and derive its transient distribution conditional on its history.

Keywords: Gaussian fluid queue, Gaussian Markov process, multivariate Brownian bridge, multivariate Brownian motion, regulated Brownian motion, stationary increments

1. INTRODUCTION

Multivariate *Brownian motion (BM)* is a Gaussian Markov process (GMP) with a mean and initial state of zero, increments that are stationary and independent, and a covariance kernel (a.k.a. covariance function) that is smooth and symmetric. Multivariate BM has been widely applied because of its analytical tractability and its ability to model variances of its components' individual states and correlations between their concurrent states. Its applications are limited, however, by its independent increments, which preclude its use for modeling nonzero correlations between its increments over nonoverlapping time intervals (i.e., nonzero autocorrelations).

In this paper, we generalize multivariate BM by dropping the property of independent increments while retaining the other properties of BM listed above. We obtain the new result that any GMP with a mean and initial state of zero, stationary increments, and a symmetric smooth kernel has continuous sample paths with probability one (w.p.1) and a kernel of a particular parametric form. We refer to such a process as a ψ-GMP (pronounced SI-GMP) as a mnemonic for a *Stationary-Increment GMP*. Multivariate BM is a special case. The kernel of an m-variate ψ-GMP depends on a total of $m(m+1)$ potentially unique scaler parameters—twice the number of parameters as the kernel of m-variate BM. Hence, the kernel of a univariate ψ-GMP depends on two parameters—one more than univariate

BM. The extra parameters enable a ψ-GMP to model autocorrelations that may be positive, negative, or zero. We show how a ψ-GMP may be represented as the solution to a linear Stochastic Differential Equation (SDE) driven by standard BM.

To provide a method for fitting a ψ-GMP to measurements, we derive the maximum-likelihood, sufficient, unbiased estimators of the kernel parameters of a ψ-GMP in closed form. That result is also new and a rare example for which closed-form maximum-likelihood estimators have been derived for a discretely sampled solution to a multivariate SDE. The estimators have the desirable property of applying for samples that are arbitrarily spaced in time.

A prominent application of BM is the analysis of queues or inventories. The random function of time describing the cumulative difference between the supply and demand for a queue's content is called the queue's net input process, and the length of a single-server queue at each time epoch is determined by the history of the net input process up to that time. Under the assumptions that the queue has an infinite buffer and a net input process that is a univariate BM with drift, the queue length process is (univariate) Reflected BM (RBM) with a well-known transient distribution. For the generalization in which the net input process is a univariate ψ-GMP with drift, we derive the transient distribution of the queue length process conditional on both its own state and the history of its net input process. The result shows how the conditional queue length distribution reflects the positive or negative autocorrelation structure of a ψ-GMP. Our derivation of the conditional queue length distribution is useful for forecasting.

We extend that last result to the case where the queue length distribution is conditioned on the history of a multivariate process, of which the queue's net input process is one component. The generalization enables the modeling of a queue that depends directly on its net input process, but indirectly on other components of a multivariate process on which the net input process depends. Consider, for example, a queue representing the volume of water at a reservoir. Its net input process represents the difference between the cumulative supply and demand for water. That net input process is itself likely to depend on the recent history of the local rainfall process (through the supply) and the local temperature process (through the demand). If so, then conditioning on histories of the net input process, the rainfall process, and the temperature process jointly is likely to provide more relevant information for forecasting the reservoir's volume than conditioning on the history of the net input process alone. We derive the conditional queue length distribution in the closed form under the assumption that the net input process and other processes on which the net input process may depend are components of a multivariate ψ-GMP with drift.

The generalization of BM to allow dependence between increments is most useful for modeling queuing systems when their unconditional and conditional behavior can be expected to markedly differ. When a queue's net input process is a ψ-GMPs with a constant drift, we show that conditioning on its state and that of the queue length process at a given time induces a change in the subsequent drift of the net input process. The change in drift persists over the remainder of the queue's time domain. Such a process for which early history persistently influences subsequent evolution may be described as self-organizing or self-exciting. Because the conditional (induced) drift can be positive when the unconditional (original) drift was negative, the conditional queue length distribution can differ persistently and markedly from the unconditional queue length distribution. As a result, ψ-GMPs are natural models for queues with self-organizing or self-exciting arrival or service processes, and the understanding of their conditional behavior developed here is crucial for their accurate prediction. For a survey of self-organizing and self-exciting processes in applications, see Arthur et al. [4].

1.1. Related Work

Our new results characterizing ψ-GMPs are related to the literature on Gaussian processes from the mid-20th century. Properties of univariate GMPs were derived on pages 94–97 of Feller [17], pages 90 and 237–238 of Doob [15], and Hida [26] and were extended to multivariate processes by Doob [14], Beautler [6], Mandrekar [36], and Adler [1]. A spectral representation of a unidimensional Gaussian process with stationary increments (a.k.a. homogeneous increments) was obtained on pages 551–559 of Doob [15] and sample path properties of multidimensional Gaussian processes with stationary increments were derived on pages 198–215 of Adler [1]. The modeling assumption of stationary increments is common in the study of self-similar Gaussian processes; see, for example, Mandelbrot and van Ness [35]. Nevertheless, the combination of characteristics defining a ψ-GMP appears not to have been studied previously.

Any stationary Gaussian process has stationary increments; see, for example, Section 18 on pages 86–88 of Yaglom [49]. The prior literature, therefore, comes closest to considering the assumptions for a ψ-GMP in its study of stationary GMPs on the half-line. Stationary GMPs and ψ-GMPs differ, however, in their assumptions about the state of the process at time zero: the initial state must be normally distributed with a nonzero covariance matrix for a stationary GMP but constant and equal to zero for a ψ-GMP. For a Gaussian process to be stationary, the parameters of its initial covariance matrix must agree with other properties of the process, for example, with its infinitesimal parameters when the GMP is defined by an SDE. It is, therefore, not possible to obtain a ψ-GMP as the limit of a stationary Gaussian Markov property as its initial covariance matrix approaches zero. Like a ψ-GMP, a stationary GMP has a parametric covariance kernel as described by Theorem 6.7 on page 357 of Kratzas and Shreve [31], but its functional form differs from that of ψ-GMP because of its different initial conditions.

Aside from BM itself, the best-known example of a ψ-GMP is the Brownian Bridge (BB). The multivariate BB can be obtained by conditioning multivariate BM on its state at a future time. The BB has been applied in the univariate case as a model for excursions with known endpoints; see, for example, Ball and Torous [5] for its use in modeling returns on bonds and Horne [29] and Fischer *et al.* [21] in modeling migrations. The covariance kernel for a multivariate BB describes the propensity of the different component processes to move in tandem.

Challenges in obtaining closed-form likelihood estimators for discretely sampled solutions to SDEs are described in Aït-Sahalia [2,3], but the Gaussian distribution of a multivariate ψ-GMP makes it more amenable to analysis than the solution to a general SDE. In deriving maximum-likelihood estimators for ψ-GMPs, we first obtain the maximum-likelihood estimator for a multivariate BB. The unique solution to the likelihood equation for a broader category of multivariate Gaussian distributions covering that case was previously obtained by Dutilleul [16]. We provide a different proof covering the new results that the solution for a BB is indeed the global maximum of the likelihood equation and possesses a closed-form spectral decomposition. Immediately following that proof, we further discuss how it relates to the treatment in Dutilleul [16].

Multivariate RBM is the heavy traffic limit for a network of queues as the traffic intensities for the queues approaches unity; see Whitt [45], Reiman [41], Harrison and Reiman [25], and Dai [8] for background. The special case of univariate RBM with a single reflecting barrier at zero is the heavy traffic limit of a single-server queue with an infinite buffer. The transient and steady-state distributions of univariate RBM, in that case, are available in closed form; see pages 14–15 of Harrison [24]. The steady-state distribution depends on two parameters, one of which is determined by the queue's traffic intensity. Iglehart and Whitt

[30] and Fendick *et al.* [19,20] showed how the single remaining parameter of univariate RBM can model complex net input processes (including those with autocorrelations) for a queue in the heavy traffic limit. Queues for which increments in the net input process are positively or negatively autocorrelated appear frequently in applications. An example where the net input process has positively correlated increments is the model with bursty arrivals studied in Fendick *et al.* [20] and interpreted in (48)–(51) on page 181 of Fendick and Whitt [18]. An example where the net input process has negatively correlated increments is the queue with periodic arrivals studied by Hajek [22].

As a queue's traffic intensity decreases from unity, its queue length distribution generally depends less on the autocorrelation structure of the net input process over long time scales and more on its variability over short time scales. Because the RBM model for a single-server queue has only a single parameter that can be used to model traffic characteristics other than the traffic intensity, it cannot capture such effects. That is one reason why heavy traffic limit theorems typically overstate or understate the impact of autocorrelations when applied alone as approximations for queues at moderate traffic levels. To address that limitation, Whitt [44] applied heavy traffic approximations in combination with limit theorems for other traffic regimes through heuristics that interpolate between the regimes; see also Sriram and Whitt [43], Fendick and Whitt [18], and Fendick *et al.* [20] for related heuristics. The last of those references demonstrates that, even then, obtaining accurate approximations for moderate traffic levels is challenging for some applications because the transition between regimes can be abrupt.

A queue driven by a ψ-GMP is a generalization of RBM. It therefore may be regarded as a refinement to a heavy traffic approximation, but its additional degree of freedom enables it to model the variability and autocorrelation structure of the net input process separately. We derive its transient queue length distribution conditional on the history of the net input process, a result without a meaningful analogy for models based on RBM alone (because of BM's independent increments). The unconditional transient queue length distribution is the special case of our result in which the history length is set equal to zero. We show that the sensitivity of the conditional and unconditional queue length distributions to autocorrelations decreases as the traffic intensity increases or decreases from unity (or, equivalently, as the drift of the net input process increases or decreases from zero). That property makes the model a better candidate than RBM for stand-alone application at moderate traffic intensities.

Recent refinements to heavy traffic asymptotics by Whitt and You [46,47] also have the potential to improve estimates of queue workloads at moderate traffic intensities. The first of those references describes steady-state mean workloads, and the second describes transient workloads for time-dependent arrival rates. Our work is similar in its philosophy of using a characterization of the covariance structure over a range of time scales. It differs in its objective of describing the effects of autocorrelations on queueing performance conditional on past history.

The queuing model studied here is an example of a Gaussian fluid queue, as defined and studied by other authors under variations on the assumptions made here. Debicki and Rolski [11] and Debicki *et al.* [12,13] derive asymptotic results for queues, assuming that the net input process is Gaussian with stationary increments, but not assuming that it is Markov. Debicki and Rolski [10] derive asymptotic results for queues, assuming that the net input process is a stationary GMP. Debicki and Mandjes [9] provide background and motivation for Gaussian fluid queueing models and describe open problems in their theory, including (i) speed of convergence to stationarity and (ii) correlation structure for the queueing process. They note that "for a general Gaussian input process .... there are no explicit expressions available for the (stationary) distribution of (the queue length) ... let alone the transient".

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

For the ψ-GMP queuing model, we obtain an explicit expression for the transient queue-length distribution, which can be used to explore the approach to stationarity, and for the conditional joint density for queue lengths at different times, which can be used to study the correlation structure of the queueing process.

1.2. Organization of Paper

The remainder of the paper is organized as follows. Section 2 states all of our results and explores their implications. Section 3 describes some examples of ψ-GMPs. Section 4 contains all proofs of the Section 2 results. The proofs about queueing results illustrate how the defining properties of a ψ-GMP are convenient for deriving tractable models. The proofs are otherwise straightforward applications of matrix theory. Their simplicity illustrates how we have preserved the tractability of BM in generalizing it.

2. STATEMENT OF THE RESULTS

In the statements that follow, we will present scalar-valued quantities in italicized type (e.g., X) and vector- or matrix-valued quantities in italicized bold type (e.g., $\boldsymbol{X}$). The transpose of the matrix or vector $\boldsymbol{a}$ will be denoted as $\boldsymbol{a}^{\mathrm{T}}$. We will say that an $\mathbb{R}^m$-valued stochastic process $\{\boldsymbol{X}(t) \equiv (X_1(t), X_2(t), \ldots, X_m(t))^{\mathrm{T}} : 0 \leq t \leq T < \infty\}$ is *Gaussian* if, for any positive integer k and any real numbers $0 < t_1 < t_2 < \cdots < t_k$, the random matrix $(\boldsymbol{X}(t_1), \boldsymbol{X}(t_2), \ldots, \boldsymbol{X}(t_k))$ of dimension $m \times k$ has a joint normal distribution. We will let $\boldsymbol{\Gamma}(t,s) \equiv E[\boldsymbol{X}(t)\boldsymbol{X}^{\mathrm{T}}(s)]$ for $0 \leq s, t \leq T$ denote the *covariance kernel* of $\boldsymbol{X}$. By the definition of a covariance kernel,

$$\boldsymbol{\Gamma}(s,t) = (\boldsymbol{\Gamma}(t,s))^{\mathrm{T}} \quad \text{for } 0 \leq s \leq t \leq T. \tag{2.1}$$

We say that the covariance kernel in Eq. (2.1) is *symmetric* if $\boldsymbol{\Gamma}(t,s) = \boldsymbol{\Gamma}(s,t)$ for $0 \leq s \leq t \leq T$. If the distribution of $\{\boldsymbol{X}_\varepsilon(t) \equiv \boldsymbol{X}(t+\varepsilon) - \boldsymbol{X}(\varepsilon) : 0 \leq t \leq T - \varepsilon\}$ does not depend on ε for any $0 \leq \varepsilon < T$, then we will say that $\boldsymbol{X}$ has *stationary increments*. When $\boldsymbol{X}(0) = 0$, $\boldsymbol{X}$ has stationary increments if and only if $\{\boldsymbol{X}_\varepsilon(t) \equiv \boldsymbol{X}(t+\varepsilon) - \boldsymbol{X}(\varepsilon) : 0 \leq t \leq T - \varepsilon\}$has the same distribution as $\{\boldsymbol{X}(t) : 0 \leq t \leq T - \varepsilon\}$ for any $0 \leq \varepsilon < T$.

The first proposition compiles known results that we will apply.

Proposition 1: *If $\{\boldsymbol{X}(t) : 0 \leq t \leq T < \infty$ and $\boldsymbol{X}(0) = 0\}$ is a zero-mean, $\mathbb{R}^m$-valued GMP for which $\boldsymbol{\Gamma}(t,s)$ is continuous at every diagonal point $s = t$ and $\boldsymbol{\Gamma}(t,t)$ is nonsingular for $0 \leq t < T$, then*

$$\boldsymbol{\Gamma}(t,s) = \boldsymbol{f}(t)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s) \quad \text{for } 0 \leq s \leq t \leq T, \tag{2.2}$$

where $\boldsymbol{f}(t)$ is nonsingular for each t with $\boldsymbol{f}(0) = \boldsymbol{I}$ and where $\boldsymbol{h}(s)$ is symmetric positive definite for $0 \leq s < T$ with $\boldsymbol{h}(0) = 0$.

The first theorem describes how additional assumptions imply a parametric covariance kernel.

Theorem 1: *Under the conditions of Proposition 1, if (i) $\boldsymbol{X}$ has stationary increments, (ii) $\boldsymbol{\Gamma}(t,s)$ is twice continuously differentiable in both s and t, (iii) $\boldsymbol{\Gamma}(t,s) = \boldsymbol{\Gamma}(s,t)$, and* *(iv)* $\boldsymbol{\alpha} \equiv (d/dt)\boldsymbol{\Gamma}(t,t)|_{t=0}$ *is symmetric positive definite, then*

$$\boldsymbol{\Gamma}(t,s) = \begin{cases} s(\boldsymbol{\alpha} - \boldsymbol{\beta}t), & 0 \le s \le t \le T \\ t(\boldsymbol{\alpha} - \boldsymbol{\beta}s), & T \ge s \ge t \ge 0 \end{cases} \qquad \mathbf{(2.3)}$$

for some symmetric matrix $\boldsymbol{\beta}$

We paraphrase the assumptions of Theorem 1 as saying that $\boldsymbol{X}$ is a GMP with stationary increments and a smooth symmetric kernel. Assumption (iii) holds trivially in the unidimensional case.

To prove the converse of Theorem 1, we will need the following result, which we will also apply to analyze the examples of Section 3.

Lemma 1: *If* $\{\boldsymbol{X}(t) : 0 \le t \le T < \infty \quad \text{and} \quad \boldsymbol{X}(0) = 0\}$ *is an* $\mathbb{R}^m$*-valued Gaussian process with the covariance kernel in Eq. (2.3) where* $\boldsymbol{\alpha}$ *is positive definite, then* $\boldsymbol{\Gamma}(t,t)$ *is positive definite for all* $0 < t < T$*. If* $\boldsymbol{\Gamma}(T,T)$ *is positive definite, then the time interval over which* $\boldsymbol{X}$ *and the arguments of its covariance kernel in Eq. (2.3) are defined may be extended to* $[0,\tilde{T}]$ *for some* $\tilde{T} > T$*. If* $\boldsymbol{\Gamma}(T,T)$ *is positive semi-definite but not positive definite, then* $\boldsymbol{\alpha} - \boldsymbol{\beta}\mathrm{t}$ *is negative definite for any* $t > T$*; and the time domain of* $\boldsymbol{X}$ *cannot then be extended beyond* $[0,T]$*.*

The next theorem describes the converse of Theorem 1.

Theorem 2: *If* $\{\boldsymbol{X}(t) : 0 \le t \le T < \infty \quad \text{and} \quad \boldsymbol{X}(0) = 0\}$ *is an* $\mathbb{R}^m$*-valued Gaussian process for which Eq. (2.3) holds where* $\boldsymbol{\alpha}$ *is positive definite, then* $\boldsymbol{X}$ *has stationary increments and the Markov property.*

We define an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,T]$ as any zero-mean Gaussian process $\{\boldsymbol{X}(t) : 0 \le t \le T < \infty \quad \text{and} \quad \boldsymbol{X}(0) = 0\}$ with the covariance kernel given by Eq. (2.3) where $\boldsymbol{\alpha}$ is positive definite. We say that a process is an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,\infty)$ if it is an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,T]$ for all $T \ge 0$. Section 3 explores conditions under which a ψ-GMP's time domain can be extended beyond a finite interval.

Let α_{ij} and β_{ij} denote the (i,j)th elements of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. The next result follows immediately from the definition of a ψ-GMP.

Corollary 1: *If* $\boldsymbol{X}$ *is an* $\mathbb{R}^m$*- valued* $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ*-GMP on* $[0,T]$*, then its ith component* X_i *for* $i = 1,\ldots,m$ *is an* $\mathbb{R}^1$*-valued* (α_{ii},β_{ii}) ψ*-GMP on* $[0,T]$ *.*

The proof of the next corollary applies Corollary 1 to show that a well-known property of BM holds more generally for ψ-GMPs.

Corollary 2: *If* $\boldsymbol{X}$ *is an* $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ*-GMP on* $[0,T]$*, then* $\boldsymbol{X}$ *has continuous sample paths w.p.1.*

A ψ-GMP can exhibit positive or negative correlations between the concurrent states of its component processes as determined by the covariance matrix $\boldsymbol{\Gamma}(t,t)$ for each $t \in [0,T]$. A ψ-GMP also can exhibit positive or negative correlations between its increments over disjoint time intervals. Our next result characterizes such autocorrelations. In the statement of that result, let $\rho(X,Y) \equiv \mathrm{Cov}(X,Y)/\left(\sqrt{\mathrm{Var}\,X}\sqrt{\mathrm{Var}\,Y}\right)$ denote the correlation coefficient for the scalar-valued random variables X and Y. As is well-known, $-1 \le \rho(X,Y) \le 1$ whenever it exists.

Proposition 2: *If $\boldsymbol{X}$ is an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,T]$, and X_i is its ith component process, then*

$$E[(\boldsymbol{X}(t+u+s)-\boldsymbol{X}(t+u))(\boldsymbol{X}(t+s)-\boldsymbol{X}(t))^{\mathrm{T}}]=-s^2\boldsymbol{\beta}$$

and

$$\rho(X_i(t+u+s)-X_i(t+u),X_i(t+s)-X_i(t))=\frac{-s\beta_{ii}}{\alpha_{ii}-s\beta_{ii}}$$

for $t\geq 0$, $u\geq s\geq 0$, and $s+t+u\leq T$, as the result of which

$$\lim_{\beta_{ii}\to-\infty}\rho(X_i(t+u+s)-X_i(t+u),X_i(t+s)-X_i(t))=1$$

and

$$\lim_{\substack{\beta_{ii}\to\alpha_{ii}/T\\ s\to T/2}}\rho(X_i(t+u+s)-X_i(t+u),X_i(t+s)-X_i(t))=-1.$$

The correlation coefficient applied in Proposition 2 has been previously used to quantify autocorrelations in empirical data; see Cont [7], for example. It describes dependence between increments of X_i over nonoverlapping intervals. The intervals are required not to overlap, because overlap creates a positive contribution to the correlation coefficient that would make the result hard to interpret. (If, in an extreme example, the intervals were allowed to exactly coincide, the resulting correlation coefficient would be a positive value equal to one regardless of other properties of the process.) Proposition 2 shows that ψ-GMPs can exhibit the maximum possible range of positive or negative autocorrelation coefficients for their component processes.

A *Wiener process or standard BM* $\{\boldsymbol{W}(t):t\geq 0\}$ is an $(\boldsymbol{I},0)$ ψ-GMP, as follows from the definition on page 184 of Adler [1]. A solution to a linear SDE driven by a Wiener process is a Markov process and is also Gaussian when the initial state is constant and coefficients are deterministic functions of time. The next result provides an interpretation of a multivariate ψ-GMP as a solution to such a linear SDE.

Theorem 3: *The $\mathbb{R}^m$-valued process $\{\boldsymbol{X}(t):0\leq t\leq T\}$ is an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,T]$ if it solves*

$$d\boldsymbol{X}(t)=\boldsymbol{\mu}(t)\boldsymbol{X}(t)dt+\boldsymbol{\sigma}(t)d\boldsymbol{W}(t)\quad\text{for } 0\leq t<T, \tag{2.4}$$

$$\boldsymbol{X}(t)=\ 0,$$

where

$$\boldsymbol{\mu}(t)=-\ \boldsymbol{\beta}(\boldsymbol{\alpha}-\boldsymbol{\beta}t)^{-1}\quad\text{for } 0\leq t<T \tag{2.5}$$

and

$$\boldsymbol{\sigma}(t)\boldsymbol{\sigma}^{\mathrm{T}}(t)=\boldsymbol{\alpha}\quad\text{for } 0\leq t<T, \tag{2.6}$$

such that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are constant symmetric matrices of dimension $m\times m$ and $\boldsymbol{\alpha}-\boldsymbol{\beta}t$ is positive definite for all $t\in[0,T)$.

According to Theorem 3, the matrix $\boldsymbol{\alpha}$ describes the instantaneous covariances between the ψ-GMP's component processes at any common point in time, and the matrix $\boldsymbol{\beta}$ describes their individual propensities to regress toward or away from zero.

For the special case in which $\boldsymbol{\beta}=0$, Eqs. (2.4)–(2.6) define a multivariate BM with instantaneous covariance matrix $\boldsymbol{\alpha}$. In that case, the multivariate BM is an $(\boldsymbol{\alpha},0)$ ψ-GMP.

On the other hands, Proposition 2 implies that $\boldsymbol{\beta} = 0$ for a ψ-GMP with independent increments. The next corollary immediately follows and justifies our description of a ψ-GMP as multivariate BM without the independent increments.

COROLLARY 3: *A ψ-GMP with independent increments is a multivariate BM with a positive definite instantaneous covariance matrix.*

In the setting of Corollary 1, if both $\beta_{ii} > 0$ and $T = \alpha_{ii}/\beta_{ii}$, then $E[X_i(T)^2] = 0$; so that $X_i(T) = 0$ w.p.1. We can, therefore, interpret $\{X_i(t) : 0 \leq t \leq \alpha_{ii}/\beta_{ii}\}$ as a univariate Gaussian bridge on $[0, \alpha_{ii}/\beta_{ii}]$. It is, in fact, a scaled BB as follows from Definition 6.12 at the top of page 360 of Karatzas and Shreve [31] and the discussion that immediately follows that definition.

If $\{\boldsymbol{Z}(t) \equiv t\boldsymbol{\omega} + \boldsymbol{X}(t) : 0 \leq t \leq T\}$ where $\boldsymbol{\omega}$ is a constant vector and $\boldsymbol{X}$ is an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP on $[0, T]$, then we will call $\boldsymbol{Z}$ an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP with drift $\boldsymbol{\omega}$ on $[0, T]$. A multivariate Gaussian bridge is obtained by conditioning such a process on a future state as described next.

LEMMA 2: *If $\{\boldsymbol{Z}(t) : 0 \leq t \leq T\}$ is an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP with drift $\boldsymbol{\omega}$ and if $\{\boldsymbol{B}(t) : 0 \leq t \leq u \leq T\}$ is the process obtained by conditioning $\{\boldsymbol{Z}(t) : 0 \leq t \leq u\}$ on $\boldsymbol{Z}(u)$, then $\boldsymbol{B}$ is an $(\boldsymbol{\alpha}, u^{-1}\boldsymbol{\alpha})$ ψ-GMP on $[0, u]$ with drift $u^{-1}\boldsymbol{Z}(u)$.*

Lemma 2 implies that $E[(\boldsymbol{B}(u) - \boldsymbol{Z}(u))(\boldsymbol{B}(u) - \boldsymbol{Z}(u))^{\mathrm{T}}] = 0$, as it must since it was assumed that $\boldsymbol{B}(u) = \boldsymbol{Z}(u)$. A surprising consequence of Lemma 2 is that $\boldsymbol{B}$ depends on neither the drift vector $\boldsymbol{\omega}$ nor the parameter matrix $\boldsymbol{\beta}$ of $\boldsymbol{Z}$. The process $\boldsymbol{B}$ is, in fact, a multivariate Brownian bridge because the same process $\boldsymbol{B}$ is obtained for the general case in which the original process $\boldsymbol{Z}$ is an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP with drift $\boldsymbol{\omega}$ as for the particular case in which the original process is an $(\boldsymbol{\alpha}, 0)$ ψ-GMP with drift 0.

The next lemma describes the maximum-likelihood estimator for the parameter matrix $\boldsymbol{\alpha}$ of the process $\boldsymbol{B}$ from Lemma 2, as obtained from a single realization of $\boldsymbol{B}$.

LEMMA 3: *If $\boldsymbol{B}$ is an $\mathbb{R}^m$-valued $(\boldsymbol{\alpha}, t_n^{-1}\boldsymbol{\alpha})$ ψ-GMP with drift $t_n^{-1}\boldsymbol{x}_n$ on $[0, t_n]$ and if sample vectors $\boldsymbol{B}(t_i) = \boldsymbol{x}_i$ are known for $0 < t_1 < t_2 < \cdots < t_{n-1} < t_n$ and some $n \geq m+1$, then*

$$\hat{\boldsymbol{\alpha}} = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{(t_{i+1}\boldsymbol{x}_i - t_i\boldsymbol{x}_{i+1})(t_{i+1}\boldsymbol{x}_i - t_i\boldsymbol{x}_{i+1})^{\mathrm{T}}}{t_i t_{i+1}(t_{i+1} - t_i)} \tag{2.7}$$

is both a sufficient estimator and the unique maximum-likelihood estimator of $\boldsymbol{\alpha}$.

Theorem 4, stated next, describes estimators for both parameter matrices of a ψ-GMP as also obtained from a single realization.

THEOREM 4: *If $\boldsymbol{X}$ is an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP on $[0, t_n]$ and if sample vectors $\boldsymbol{X}(t_i) = \boldsymbol{x}_i$ are known for $0 < t_1 < t_2 < \cdots < t_n$, then $\hat{\boldsymbol{\alpha}}$ in Eq. (2.7) and*

$$\hat{\boldsymbol{\beta}} = \frac{\hat{\boldsymbol{\alpha}}}{t_n} - \frac{\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}}}{t_n^2} \tag{2.8}$$

are the unique maximum-likelihood, jointly sufficient, unbiased estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The next lemma is similar to Lemma 2, but its result is obtained by conditioning on a past rather than the future state.

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

LEMMA 4: *If* $\{\boldsymbol{Z}(t) : 0 \leq t \leq T\}$ *is an* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ*-GMP with drift* $\boldsymbol{\omega}$*, then the process obtained by conditioning* $\{\boldsymbol{Z}^{(s)}(t) \equiv \boldsymbol{Z}(s+t) - \boldsymbol{Z}(s) : 0 \leq t \leq T-s\}$ *on* $\{\boldsymbol{Z}(v) : 0 \leq v \leq s\}$ *for fixed* $0 < s < T$ *is an* $(\boldsymbol{\alpha}, \boldsymbol{\beta_s})$ ψ*-GMP with drift* $\boldsymbol{\omega}_s$ *on* $[0, T-s]$*, where*

$$\boldsymbol{\omega}_s \equiv \boldsymbol{\omega} - \boldsymbol{\beta}(\boldsymbol{\alpha} - \boldsymbol{\beta} s)^{-1}(\boldsymbol{Z}(s) - s\boldsymbol{\omega}) \quad \text{and} \quad \boldsymbol{\beta}_s \equiv \boldsymbol{\beta}(\boldsymbol{\alpha} - \boldsymbol{\beta} s)^{-1}\boldsymbol{\alpha}\,. \tag{2.9}$$

According to (2.9), the change in drift from $\boldsymbol{\omega}$ to $\boldsymbol{\omega}_s$ can be arbitrarily large and can even result in a change in sign.

The last two theorems apply Lemmas 2 and 4 to characterize a queuing process. Given any $\mathbb{R}^1$-valued continuous process $\{Z(t) :\ 0 \leq t \leq T \ \text{ and } \ Z(0) = 0\}$, let $\{Q(t) :\ 0 \leq t \leq T\}$ denote the solution to

$$Q(t) = Q(0) + Z(t) + L(t) \geq 0 \quad \text{for } 0 \leq t \leq T, \tag{2.10}$$

where $Q(0)$ is a constant and $L(\cdot)$ is a nondecreasing continuous function with the properties that (i) $L(0) = 0$ and (ii) $L(\cdot)$ increases only when $Q(\cdot) = 0$ such that $Q(t) \geq 0$ for all $0 \leq t \leq T$. At time t, $Q(t)$ has the interpretation as the content of a queue, and $Z(t)$ as the queue's cumulative net input up to that time (interpreted as the difference between the cumulative supply and the cumulative demand for the queue's content). The nondecreasing process $L(t)$ then has the interpretation as the cumulative demand for the queue's content that is unsatisfied over the interval $[0, t]$. The functional defined by Eq. (2.10) mapping $\{Z(s) :\ 0 \leq s \leq t\}$ to $Q(t)$ for each t is variously known as the *univariate reflection map*, the *one-sided regulator*, and the *Skorokhod map;* see Section 5.2.2 and Chapter 14 of Whitt [45] or Section 2 of Harrison [24] for background.

THEOREM 5: *If* Q *is defined by Eq. (2.10) where* Z *is an* $\mathbb{R}^1$*- valued* (α, β) ψ*-GMP with drift* ω *on* $[0, T]$*, then*

$$\begin{aligned} &P(Q(s+u) \leq q_{s+u} | Z(s) = z_s, Q(s) = q_s) \\ &= \frac{1}{2}\left(1 - e^{\frac{-2q_{s+u}(\beta_s q_{s+u} - \alpha\omega_s)}{\alpha^2}} - erf\left(\frac{-q_{s+u} + q_s + \omega_s u}{\sqrt{2\mathrm{u}(\alpha - \beta_s \mathrm{u})}}\right)\right. \\ &\quad \left. + e^{\frac{-2q_{s+u}(\beta_s q_{s+u} - \alpha\omega_s)}{\alpha^2}} erf\left(\frac{\alpha(q_{s+u} + q_s) - (2\beta_s q_{s+u} - \alpha\omega_s)\mathrm{u}}{\alpha\sqrt{2\mathrm{u}(\alpha - \beta_s \mathrm{u})}}\right)\right) \end{aligned} \tag{2.11}$$

for $q_{s+u} \geq 0$ *and* $0 < s \leq s + u < T$*, where*

$$\omega_s \equiv \frac{\alpha\omega - \beta z_s}{\alpha - \beta s}, \quad \beta_s \equiv \frac{\alpha\beta}{\alpha - \beta s}\,, \tag{2.12}$$

and

$$erf(x) \equiv 2\pi^{-1/2} \int_0^x \exp(-t^2)dt.$$

An example of a process obtained by applying the reflection map in Eq. (2.10) to a ψ-GMP with drift is the heavy traffic limit for a sequences of queues in which a fixed number of arrivals occur with mutually independent, uniformly distributed arrival times over an interval and in which service times are deterministic. In that case, Theorem 1 of Iglehart and Whitt [30] and the results on page 365 there show that the limit for the net input process is a $(1, 1)$ ψ-GMP with constant drift on the time domain $[0, 1]$, that is, a standard BB with drift. Hajek [22] provided an alternative derivation of that result to model a queue with periodic arrivals and obtained the unconditional queue length distribution at the upper limit of the time domain for the reflected diffusion. The proof of Theorem 5 uses Hajek's result. The result in Theorem 5 generalizes Hajek's result to describe the conditional distribution over the whole time domain and the unconditional transient distribution over the whole time domain as a limiting case. Whereas the net input process in that example has a negative autocorrelation structure ($\beta > 0$), Theorem 5 applies as well for net input processes with positive autocorrelations ($\beta < 0$) or no autocorrelations ($\beta = 0$).

The distribution in Eq. (2.11) depends on the conditioned event that $Z(s) = z_s$ through the parameter ω_s as defined in Eq. (2.12). That dependence is strongest when $\omega \approx 0$. The unconditional transient distribution $P(Q(u) \leq q_u)$ generalizing the transient distribution for RBM from page 49 of Harrison [24] is obtained as the limit of Eq. (2.12) as $s \to 0$ in which case $z_s \to 0$, $q_s \to Q(0)$, $\omega_s \to \omega$, and $\beta_s \to \beta$. In that case, the dependence on autocorrelations through the parameter β is also strongest when $\omega \approx 0$. In general, the parameter ω_s may be interpreted as the conditional (or induced) drift of the net input process Z on $[s, T]$ given $Z(s)$. The conditional drift ω_s can be positive even when the unconditional drift ω is negative.

In the statement of the following corollaries to Theorem 5, $\delta(\cdot)$ denotes the Dirac delta function, and 1_A denotes the indicator function equal to 1 on A and 0 zero elsewhere.

Corollary 4: *Under the conditions of Theorem 5,*

$$
\begin{aligned}
&P(Q(s) \in dq_s, Q(s+u) \in dq_{s+u} | Z(s) = z_s) \\
&\quad = P(Q(s+u) \in dq_{s+u} \mid Z(s) = z_s, Q(s) = q_s) P(Q(s) \in dq_s | Z(s) = z_s)
\end{aligned} \tag{2.13}
$$

and

$$
\begin{aligned}
&P(Z(s) \in dz_s, Q(s) \in dq_s, Q(s+u) \in dq_{s+u}) \\
&\quad = P(Q(s+u) \in dq_{s+u} | Z(s) = z_s, Q(s) = q_s) P(Q(s) \in dq_s | Z(s) = z_s) P(Z(s) \in dz_s),
\end{aligned} \tag{2.14}
$$

where

$$
\begin{aligned}
&P(Q(s+u) \in dq_{s+u} | Z(s) = z_s, Q(s) = q_s) \\
&= \frac{1}{2} \left(\frac{\sqrt{2}(\alpha - 2u\beta_s) e^{-\frac{(\alpha(q_{s+u}+q_s)-(2\beta_s q_{s+u}-\alpha\omega_s)\mathrm{u})^2}{2\alpha^2 \mathrm{u}(\alpha-\beta_s \mathrm{u})} - \frac{2q_{s+u}(\beta_s q_{s+u}-\alpha\omega_s)}{\alpha^2}}}{\alpha\sqrt{\pi u(\alpha - \beta_s \mathrm{u})}} \right. \\
&\quad + \frac{\sqrt{2} e^{-\frac{(-q_{s+u}+q_s+\omega_s u)^2}{2\mathrm{u}(\alpha-\beta_s \mathrm{u})}}}{\sqrt{\pi u(\alpha - \beta_s \mathrm{u})}} - \frac{4\beta_s q_{s+u} - 2\alpha\omega_s}{\alpha^2} e^{-\frac{2q_{s+u}(\beta_s q_{s+u} - \alpha\omega_s)}{\alpha^2}} \\
&\quad \times erf\left(\frac{\alpha(q_{s+u} + q_s) - (2\beta_s q_{s+u} - \alpha\omega_s)\mathrm{u}}{\alpha\sqrt{2u(\alpha - \beta_s \mathrm{u})}} \right) + \frac{4\beta_s q_{s+u} - 2\alpha\omega_s}{\alpha^2} \\
&\quad \left. \times\, e^{-\frac{2q(\beta_s q_{s+u} - \alpha\omega_s)}{\alpha^2}} \right) dq_{s+u},
\end{aligned} \tag{2.15}
$$

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

$$P(Q(s)\in dq_s|Z(s)=z_s) \\ = \left(\left(1-e^{\frac{-2q_s(q_s-z_s)}{\alpha s}}\right)(\delta(q_s-Q(0)-z_s)1_{\{q_s\geq 0\}}+1_{\{q_s\geq Q(0)+z_s\}}\delta(q_s)) \right. \\ \left. +\frac{(4q_s-2z_s)}{\alpha s}e^{\frac{-2q_s(q_s-z_s)}{\alpha s}}1_{\{q_s\geq Q(0)+z_s\}}1_{\{q_s\geq 0\}}\right)dq_s, \quad \textbf{(2.16)}$$

and

$$P(Z(s)\in dz_s)=\exp(-(z_s-\omega s)^2/(2s(\alpha-\beta s)))dz_s/\sqrt{2\pi s(\alpha-\beta s)}, \quad \textbf{(2.17)}$$

where ω_s and β_s are defined in Eq. (2.12).

The last theorem generalizes Theorem 5 for the case in which the net input process to the queue is one component of a multivariate ψ-GMP with drift. In its statement, $[\boldsymbol{v}]_i$ denotes the ith component of the vector $\boldsymbol{v}$**,** and $[\boldsymbol{a}]_{ii}$ denotes the (i,i)th element of the matrix $\boldsymbol{a}$**.**

Theorem 6: *If Q is defined by Eq. (2.10) where $Z\equiv Z_i$ is the ith component of an $\mathbb{R}^m$- valued process $\boldsymbol{Z}$ that is an $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP with drift $\boldsymbol{\omega}$ on $[0,T]$, then $P(Q(s+u)\leq q_{s+u}\mid \boldsymbol{Z}(s)=\boldsymbol{z}_s, Q(s)=q_s)$ is given by the right-hand side of Eq. (2.11) for $q_{s+u}\geq 0$ and $0<s\leq s+u<T$, where*

$$\omega_s\equiv[\boldsymbol{\omega}-\boldsymbol{\beta}(\boldsymbol{\alpha}-\boldsymbol{\beta}s)^{-1}(\boldsymbol{z}_s-s\boldsymbol{\omega})]_i,\quad \alpha\equiv[\boldsymbol{\alpha}]_{ii}=\alpha_{ii},\quad \beta_s\equiv[\boldsymbol{\beta}(\boldsymbol{\alpha}-\boldsymbol{\beta}s)^{-1}\boldsymbol{\alpha}]_{ii}. \quad \textbf{(2.18)}$$

3. EXAMPLES

The following examples describe the time domain over which a ψ-GMP can be defined and reveal practical considerations in applying its maximum-likelihood estimators from Theorem 4.

Example 1: *If X is a univariate (α,β) ψ-GMP and $\beta\leq 0$, then X can be defined as a ψ-GMP on $[0,T]$ for any $T\geq 0$ since* $\operatorname{Var} X(t)=t(\alpha-\beta t)\geq 0$ *for all $t\geq 0$. In that case, X can be defined as an (α,β) ψ-GMP on $[0,\infty)$.*

Example 2: *If X is a univariate (α,β) ψ-GMP and $\beta>0$, then X can be defined as a ψ-GMP on $[0,T]$ for $T\leq\alpha/\beta$, but not for larger values of T since* $\operatorname{Var} X(t)=t(\alpha-\beta t)\geq 0$ *only for $0\leq t\leq\alpha/\beta$.*

Example 3: *Suppose that $\boldsymbol{X}$ is an $\mathbb{R}^m$- valued $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,T]$ where $\boldsymbol{\beta}=\boldsymbol{\alpha}/T-\boldsymbol{\gamma}/T^2$ when $\boldsymbol{\gamma}$ is positive semi-definite but not positive definite. Then, $\boldsymbol{\Gamma}(T,T)=T(\boldsymbol{\alpha}-\boldsymbol{\beta}T)=\boldsymbol{\gamma}$. According to Lemma 1, $\boldsymbol{\Gamma}(t,t)$ is then positive definite for $0<t<T$, but $\boldsymbol{\alpha}-\boldsymbol{\beta}t$ is negative definite for $t>T$. Therefore, $\boldsymbol{X}$ cannot be defined as a multivariate $(\boldsymbol{\alpha},\boldsymbol{\beta})$ ψ-GMP on $[0,\tilde{T}]$ for any $\tilde{T}>T$.*

Example 4: *A special case of Example 3 is obtained when $\beta_{ii}>0$ for some i. Then, Corollary 1 implies that* $\operatorname{Var} X_i(t)=t(\alpha_{ii}-\beta_{ii}t)\geq 0$ *only for $t\leq\alpha_{ii}/\beta_{ii}$. The quantity $\alpha_{ii}-\beta_{ii}t$ will become progressively more negative with increasing $t>\alpha_{ii}/\beta_{ii}$. Therefore, the multivariate process $\boldsymbol{X}$ (of which X_i is one component) can be defined as a ψ-GMP on $[0,T]$ only for $T\leq\alpha_{ii}/\beta_{ii}$.*

Example 5: *A special case of Example 3 where* $m \geq 2$ *and* $\operatorname{Var} X_i(T) > 0$ *for all* $i = 1, \ldots, m$ *is where* $\boldsymbol{\gamma} = \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}$ *for some vector* $\boldsymbol{x}$ *of dimension* $m \times 1$ *with nonzero elements. Then,* $\operatorname{Var} X_i(T) > 0$ *for all* $i = 1, \ldots, m$ *because the diagonal elements of* $\boldsymbol{\gamma}$ *are all nonzero, and* $\boldsymbol{\gamma}$ *is positive semi-definite but not positive definite because it has a rank of one.*

Example 6: *A special case of Example 5 of practical interest is where* $\boldsymbol{X}$ *is the* $\mathbb{R}^m$*-valued* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ*-GMP that results from assigning* $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ *and* $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ *using the maximum-likelihood estimators from Eqs. (2.7) and (2.8). In that case,* $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\alpha}}/t_n) - (\hat{\gamma}/t_n^2)$ *where* $\hat{\gamma} \equiv \boldsymbol{\Gamma}(t_n, t_n) = \boldsymbol{x}_n \boldsymbol{x}_n^{\mathrm{T}}$ *is a rank-one matrix. (The matrix* $\boldsymbol{\Gamma}(t_n, t_n)$ *has rank one because the maximum-likelihood estimators were obtained using a single realization of the* ψ*-GMP.) Since* $\boldsymbol{\Gamma}(t_n, t_n)$ *is then positive semi-definite but not positive definite, the process* $\boldsymbol{X}$ *is an* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ*-GMP on* $[0, t_n]$ *but not on* $[0, \tilde{T}]$ *for any* $\tilde{T} > t_n$ *(by Lemma 1).*

Example 7: *In the setting of Example 6, it may be desirable to modify the fitted* ψ*-GMP so that it can be applied over a longer interval than the interval* $[0, t_n]$ *spanned by historic samples. One approach for doing so is to let* $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ *as in Example 6 but to let* $\boldsymbol{\beta} = \hat{\hat{\boldsymbol{\beta}}} \equiv (\hat{\boldsymbol{\alpha}}/t_n) - (\hat{\hat{\gamma}}/t_n^2)$*, where* $\hat{\hat{\gamma}}$ *is a weighted sum of* $\hat{\gamma}$ *from Example 6 and a positive definite matrix with the same diagonal elements as* $\hat{\boldsymbol{\gamma}}$*. The matrix* $\hat{\hat{\gamma}}$ *is then called a shrinkage estimator. With that modification to* $\boldsymbol{\beta}$*, the covariance matrix* $\boldsymbol{\Gamma}(t_n, t_n) = \hat{\hat{\gamma}}$ *is positive definite, and Lemma 1 implies that the process with the modified parameters is an* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ*-GMP on* $[0, \tilde{T}]$ *for some* $\tilde{T} > t_n$*. Because the modification of parameters leaves the diagonal elements of both* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\beta}$ *unchanged, it leaves the autocorrelation structure unchanged for each of the component processes* X_i *individually (by Corollary 1). Methods for obtaining a shrinkage estimator of a covariance matrix are described in Ledoit and Wolf [33], Schäfer and Strimmer [42], and reference cited therein.*

Example 8: *In the setting of Example 6, it may be desirable to modify the fitted* ψ*-GMP so that it can be applied over a particular interval* $[0, \tilde{T}]$ *where* $\tilde{T} > t_n$ *is a given value. A method for doing so is as follows: first, let* $\tilde{\boldsymbol{\gamma}} = \tilde{T}(\hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}\tilde{T})$*, which is negative definite by Lemma 1 because* $\boldsymbol{\Gamma}(t_n, t_n)$ *is not positive definite when the parameters* $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ *and* $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}$ *are used; second, let* $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}$ *as before but let* $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}} \equiv \hat{\boldsymbol{\alpha}}/\tilde{T} - \tilde{\boldsymbol{\gamma}}/\tilde{T}^2$ *where* $\tilde{\boldsymbol{\gamma}}$ *is the closest symmetric positive semi-definite matrix to* $\tilde{\boldsymbol{\gamma}}$ *(in some norm). In that case,* $\boldsymbol{\Gamma}(\tilde{T}, \tilde{T}) = \tilde{T}(\boldsymbol{\alpha} - \boldsymbol{\beta}\tilde{T}) = \tilde{\boldsymbol{\gamma}}$*, and Lemma 1 implies that the process with the modified parameters is a* ψ*-GMP on* $[0, \tilde{T}]$*. Methods for obtaining the closest symmetric positive definite matrix (using the Frobenius norm) are described in Higham [27,28].*

Example 9: *If* X *is a univariate* (α, β) ψ*-GMP with* $\alpha = \hat{\alpha}$*, and* $\beta = \hat{\beta}$ *when* $\hat{\alpha}$ *and* $\hat{\beta}$ *are the maximum-likelihood estimates from Eqs. (2.7) and (2.8), then* $\Gamma(t_n, t_n) = x_n^2$*, so that* X *is not restricted to* $[0, t_n]$ *unless* $x_n = 0$.

4. PROOFS OF THE RESULTS

In the statements that follows, $\boldsymbol{a}^{-\mathrm{T}} \equiv (\boldsymbol{a}^{-1})^{\mathrm{T}} = (\boldsymbol{a}^{\mathrm{T}})^{-1}$ will denote the transpose of the inverse of the nonsingular square matrix $\boldsymbol{a}$. In addition, $\boldsymbol{b}^{-}$ will denote any generalized inverse of the matrix $\boldsymbol{b}$ as defined in Section 3.5 of Petersen and Pederson [37], and $\boldsymbol{b}^{+}$ will denote the Moore–Penrose pseudoinverse of $\boldsymbol{b}$ as defined in Section 3.6 of Petersen and

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

Pederson [37]. A process $\{\boldsymbol{X}(t) : 0 \le t \le T\}$ is continuous in quadratic mean (C.Q.M.) if

$$E[\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{X}(t)] < \infty \quad \text{and} \quad \lim_{s\uparrow t} E[(\boldsymbol{X}(t) - \boldsymbol{X}(s))^{\mathrm{T}}(\boldsymbol{X}(t) - \boldsymbol{X}(s))] = 0 \quad \text{for } 0 \le s \le t \le T.$$

The C.Q.M. process $\boldsymbol{X}$ with covariance kernel $\boldsymbol{\Gamma}(t,s) \equiv E[\boldsymbol{X}(t)\boldsymbol{X}^{\mathrm{T}}(s)]$ has the wide-sense Markov property if

$$\mathbf{A}(u,s) = \mathbf{A}(u,t)\mathbf{A}(t,s) \quad \text{for } 0 \le s \le t \le u \le T, \tag{4.1}$$

where

$$\mathbf{A}(t,s) = \boldsymbol{\Gamma}(t,s)(\boldsymbol{\Gamma}(s,s))^{+} \quad \text{for } 0 \le s \le t \le T \tag{4.2}$$

is the *transition matrix function* for the process; see Theorem 2 of Beutler [6] for equivalent statements of the wide-sense Markov property. For a Gaussian process, it is well-known that the wide-sense Markov property is equivalent to the (strict-sense) Markov property; see page 90 of Doob [15] for a proof in the univariate case that extends essentially unchanged for the multivariate case.

Proof: Since $\boldsymbol{\Gamma}(t,s)$ is continuous at every diagonal point $s = t$, $\boldsymbol{X}$ is everywhere C.Q.M. as follows from Theorem 2.2.1 on page 26 of Adler [1]. Furthermore, $\boldsymbol{X}$ is a wide-sense Markov process because it is a Markov process and also Gaussian. ■

For any C.Q.M wide-sense Markov process $\{\boldsymbol{X}(t) : 0 \le t \le T < \infty \quad \text{and} \quad \boldsymbol{X}(0) = 0\}$ for which $\boldsymbol{\Gamma}(s,s)$ is nonsingular for each $0 < s < T$, Theorem 3.1 of Mandrekar [36] implies that $\boldsymbol{X}(t) = \boldsymbol{f}(t)\boldsymbol{U}(t)$ for $0 \le t \le T$, where $\{\boldsymbol{U}(t) : 0 \le t \le T \quad \text{and} \quad \boldsymbol{U}(0) = 0\}$ is another $\mathbb{R}^m$- valued wide-sense Markov process satisfying $E[\boldsymbol{U}(t)|\boldsymbol{U}(s)] = \boldsymbol{U}(s)$ for $0 \le s \le t \le T$ and where $\boldsymbol{f}(t)$ is a nonsingular $m \times m$ matrix for each t. Since $\boldsymbol{\Gamma}(s,s) \equiv E[\boldsymbol{X}(s)\boldsymbol{X}^{\mathrm{T}}(s)] = \boldsymbol{f}(s)E[\boldsymbol{U}(s)\boldsymbol{U}^{\mathrm{T}}(s)]\boldsymbol{f}^{\mathrm{T}}(s)$ is nonsingular for $0 < s < T$, so must be $E[\boldsymbol{U}(s)\boldsymbol{U}^{\mathrm{T}}(s)]$ for $0 < s < T$. By Eqs. (2.10) and (2.12) of Beutler [6], we see that $\boldsymbol{I} = E[\boldsymbol{U}(t)\boldsymbol{U}^{\mathrm{T}}(s)](E[\boldsymbol{U}(s)\boldsymbol{U}^{\mathrm{T}}(s)])^{-1}$ for $0 < s < t \le T$ and conclude that $\boldsymbol{h}(s) \equiv E[\boldsymbol{U}(s)\boldsymbol{U}^{\mathrm{T}}(s)] = E[\boldsymbol{U}(t)\boldsymbol{U}^{\mathrm{T}}(s)]$ for $0 \le s \le t \le T$, where the equality holds trivially when $s = 0$ and when $s = t$. Therefore,

$$\boldsymbol{\Gamma}(t,s) \equiv E[\boldsymbol{X}(t)\boldsymbol{X}^{\mathrm{T}}(s)] = \boldsymbol{f}(t)E[\boldsymbol{U}(t)\boldsymbol{U}^{\mathrm{T}}(s)]\boldsymbol{f}^{\mathrm{T}}(s) = \boldsymbol{f}(t)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s) \quad \text{for } 0 \le s \le t \le T. \tag{4.3}$$

The matrix-valued function $\boldsymbol{h}(s)$ is symmetric positive semi-definite for each s because it is a covariance matrix, and $\boldsymbol{h}(0) = 0$ because $\boldsymbol{\Gamma}(0,0) = 0$. Without loss of generality, the functions $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ can be normalized so that $\boldsymbol{f}(0) = \boldsymbol{I}$ without affecting their other properties described above: if $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ have the properties described above except that $\boldsymbol{f}(0) \ne \boldsymbol{I}$, then

$$\boldsymbol{\Gamma}(t,s) = (\boldsymbol{f}(t)\boldsymbol{f}^{-1}(0))(\boldsymbol{f}(0)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(0))(\boldsymbol{f}^{-\mathrm{T}}(0)\boldsymbol{f}^{\mathrm{T}}(s)) \quad \text{for } 0 \le s \le t \le T$$

by Eq. (4.3), and the required normalizations are obtained by assigning $\boldsymbol{f}(t) \leftarrow \boldsymbol{f}(t)\boldsymbol{f}^{-1}(0)$ and $\boldsymbol{h}(s) \leftarrow \boldsymbol{f}(0)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(0)$, where the original functions are used on the right-hand sides of both assignments.

Proof: By Lemma 1 and assumption (iv),

$$\boldsymbol{f}(0) = \boldsymbol{I}, \quad \boldsymbol{h}(0) = 0, \quad \text{and} \quad \dot{\boldsymbol{h}}(0) = \boldsymbol{\alpha}. \tag{4.4}$$

■

Assumption (ii) implies that $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are each twice continuously differentiable. Let

$$\mathbf{w}(s,t) \equiv E[(\boldsymbol{X}(t+s)-\boldsymbol{X}(s))^2]. \tag{4.5}$$

By Eqs. (2.2) and (4.5),

$$\begin{aligned}\boldsymbol{w}(s,t) &= \boldsymbol{f}(t+s)\boldsymbol{h}(t+s)\boldsymbol{f}^{\mathrm{T}}(t+s) - \boldsymbol{f}(t+s)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s) \\ &\quad - \boldsymbol{f}(s)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(t+s) + \boldsymbol{f}(s)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s)\end{aligned} \tag{4.6}$$

for $0 \le s \le s+t \le T$. Assumptions (i) and (ii) then imply that $(\partial/\partial s)\boldsymbol{w}(s,t) = 0$ on $0 \le s \le s+t \le T$, so that

$$0 = \frac{\partial}{\partial s}\boldsymbol{w}(0,t) = (\boldsymbol{f}(t)\boldsymbol{h}(t)\boldsymbol{f}^{\mathrm{T}}(t))' - \boldsymbol{f}(t)\boldsymbol{\alpha} - \boldsymbol{\alpha}\boldsymbol{f}^{T}(t) + \boldsymbol{\alpha} \tag{4.7}$$

for $0 \le t \le T$ using Eqs. (4.4) and (4.6). Comparing (2.2) and (4.7), we see that

$$\boldsymbol{\Gamma}(t,t) = \left(\int_0^t \boldsymbol{f}(u)du\right)\boldsymbol{\alpha} + \boldsymbol{\alpha}\left(\int_0^t \boldsymbol{f}^{\mathrm{T}}(u)du\right) - \boldsymbol{\alpha}t, \tag{4.8}$$

so that

$$\boldsymbol{h}(t) = \boldsymbol{f}^{-1}(t)\left(\left(\int_0^t \boldsymbol{f}(u)du\right)\boldsymbol{\alpha} + \boldsymbol{\alpha}\left(\int_0^t \boldsymbol{f}^{\mathrm{T}}(u)du\right) - \boldsymbol{\alpha}t,\right)\boldsymbol{f}^{-\mathrm{T}}(t) \tag{4.9}$$

for $0 \le t < T$ by Eq. (2.2). By Eqs. (4.4) and (4.9)

$$\ddot{\boldsymbol{h}}(0) = -\dot{\boldsymbol{f}}(0)\boldsymbol{\alpha} - \boldsymbol{\alpha}\dot{\boldsymbol{f}}^{\mathrm{T}}(0). \tag{4.10}$$

Assumptions (i) and (ii) also imply that $(\partial^2/\partial s^2)\boldsymbol{w}(s,t) = 0$ for $0 \le s \le s+t < T$. Applying Eqs. (4.4), (4.6), (4.9), and (4.10), we then find that

$$\begin{aligned}0 = \frac{\partial^2}{\partial s^2}\boldsymbol{w}(0,t) &= (\dot{\boldsymbol{f}}(0) - \dot{\boldsymbol{f}}(t))\boldsymbol{\alpha} + \boldsymbol{\alpha}(\dot{\boldsymbol{f}}^{\mathrm{T}}(0) - \dot{\boldsymbol{f}}^{\mathrm{T}}(t)) \\ &\quad + \boldsymbol{f}(t)\dot{\boldsymbol{f}}(0)\boldsymbol{\alpha} - \boldsymbol{f}(t)\boldsymbol{\alpha}\dot{\boldsymbol{f}}^{\mathrm{T}}(0) - \dot{\boldsymbol{f}}(0)\boldsymbol{\alpha}\boldsymbol{f}^{\mathrm{T}}(t) + \boldsymbol{\alpha}\dot{\boldsymbol{f}}^{\mathrm{T}}(0)\boldsymbol{f}^{\mathrm{T}}(t).\end{aligned} \tag{4.11}$$

Using assumption (iii) for the first time together with Eqs. (2.1), (2.2), and (4.4), we see that

$$\boldsymbol{f}(t)\boldsymbol{\alpha} = \frac{\partial}{\partial s}\boldsymbol{\Gamma}(t,0) = \frac{\partial}{\partial s}(\boldsymbol{\Gamma}(t,0))^{T} = \boldsymbol{\alpha}\boldsymbol{f}^{\mathrm{T}}(t) \quad \text{for } 0 \le t \le T. \tag{4.12}$$

By Eqs. (4.11) and (4.12), we obtain

$$0 = \frac{\partial^2}{\partial s^2}\boldsymbol{w}(0,t) = 2(\dot{\boldsymbol{f}}(0) - \dot{\boldsymbol{f}}(t))\boldsymbol{\alpha} \quad \text{for } 0 \le t < T \tag{4.13}$$

for $0 \le t < T$. Since $\boldsymbol{\alpha}$ is nonsingular, we conclude that $\dot{\boldsymbol{f}}(t) = \dot{\boldsymbol{f}}(0)$ for all $0 \le t \le \mathrm{T}$, so that

$$\boldsymbol{f}(t) = \boldsymbol{f}(0) + \dot{\boldsymbol{f}}(0)t = \boldsymbol{I} + \dot{\boldsymbol{f}}(0)t = \boldsymbol{I} - \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}t \quad \text{for } 0 \le t \le \mathrm{T}, \tag{4.14}$$

where the second equality is obtaining using Eq. (4.4), and $\boldsymbol{\beta} \equiv -\dot{\boldsymbol{f}}(0)\boldsymbol{\alpha}$**.** The symmetry of $\boldsymbol{\beta}$ follows from Eq. (4.12).

By Eqs. (4.8), (4.12), and (4.14), we get

$$\boldsymbol{\Gamma}(s,s) = 2\left(\int_0^s \boldsymbol{f}(u)du\right)\boldsymbol{\alpha} - \boldsymbol{\alpha}s = 2\left(\int_0^s (\boldsymbol{I} + \dot{\boldsymbol{f}}(0)u)du\right)\boldsymbol{\alpha} - \boldsymbol{\alpha}s = s(\boldsymbol{\alpha} - \boldsymbol{\beta}s).$$

Then, for $0 \leq s \leq t \leq T$,

$$\begin{aligned}\boldsymbol{\Gamma}(t,s) &= \boldsymbol{f}(t)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s) = \boldsymbol{f}(t)\boldsymbol{f}^{-1}(s)\boldsymbol{f}(s)\boldsymbol{h}(s)\boldsymbol{f}^{\mathrm{T}}(s) = \boldsymbol{f}(t)\boldsymbol{f}^{-1}(s)\boldsymbol{\Gamma}(s,s)\\ &= (\boldsymbol{I} - \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}t)(\boldsymbol{I} - \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}\boldsymbol{s})^{-1}s(\boldsymbol{\alpha} - \boldsymbol{\beta}s) = \mathrm{s}(\boldsymbol{\alpha} - \boldsymbol{\beta t})(\boldsymbol{\alpha} - \boldsymbol{\beta}s)^{-1}(\boldsymbol{\alpha} - \boldsymbol{\beta}s)\\ &= s(\boldsymbol{\alpha} - \boldsymbol{\beta}t),\end{aligned}$$

where the first three equalities follow from Lemma 1, and the last three from the prior two results above. That proves Eq. (2.3) for $s \leq t$.The conclusion for $s > t$ follows from Eq. (2.1).

Proof: If $\boldsymbol{X}$ is a Gaussian process on $[0,T]$ with the covariance kernel from Eq. (2.2) where $\boldsymbol{\alpha}$ is symmetric positive definite, then $\boldsymbol{\alpha} - \boldsymbol{\beta}t$ is positive definite at $t = 0$. Since $\boldsymbol{X}$ is assumed to be a Gaussian process, $\boldsymbol{\Gamma}(t,t)$ must be positive semi-definite on $(0,T]$, so that $\boldsymbol{\alpha} - \boldsymbol{\beta}t$ must be too. Suppose that $\boldsymbol{\Gamma}(t^*,t^*)$ is positive semi-definite but not positive definite for some $t^* \in (0,T]$. By Eq. (2.3), $\boldsymbol{\beta} = \boldsymbol{\alpha}/t^* - \boldsymbol{\Gamma}(t^*,t^*)/(t^*)^2$. Consider the quantity

$$\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{\alpha} - \boldsymbol{\beta}t)\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}\left(\left(1 - \frac{t}{t^*}\right)\boldsymbol{\alpha}\right)\boldsymbol{y} + \boldsymbol{y}^{\mathrm{T}}\left(\frac{t}{(t^*)^2}\boldsymbol{\Gamma}(t^*,t^*)\right)\boldsymbol{y}, \tag{4.15}$$

where $\boldsymbol{y}$ is any nonzero vector of appropriate dimension:

- When $0 \leq t < t^*$, Eq. (4.15) is positive since $\boldsymbol{\alpha}$ is positive definite and $\boldsymbol{\Gamma}(t^*,t^*)$ is positive semi-definite. In that case, $\boldsymbol{\alpha} - \boldsymbol{\beta}t$ is positive definite.
- When $t = t^*$, $\boldsymbol{\Gamma}(t^*,t^*)$ is positive semi-definite but not positive definite by assumption, so that $\boldsymbol{\alpha} - \boldsymbol{\beta}t$ is too.
- When $t > t^*$, Eq. (4.15) is negative for some $\boldsymbol{y}$ since $(1 - (t/t^*))\boldsymbol{\alpha}$ is then negative definite and $\boldsymbol{\Gamma}(t^*,t^*)$ is positive semi-definite but not positive definite. In that case, $\boldsymbol{\alpha} - \boldsymbol{\beta}t$ is negative definite. ■

Hence, $t^* = T$. Only at $t = t^* = T$ may $\boldsymbol{\Gamma}(t,t)$ be positive semi-definite but not positive definite.

Proof: To show that Eq. (4.1) holds when $\mathbf{X}$ is a Gaussian process with the covariance kernel from Eq. (2.3) where $\boldsymbol{\alpha}$ is symmetric positive definite, we consider two cases:

If $0 < s \leq t = u \leq T$, then

$$\begin{aligned}\mathbf{A}(u,t)\mathbf{A}(t,s) &= \boldsymbol{\Gamma}(u,u)(\boldsymbol{\Gamma}(u,u))^{+}\boldsymbol{\Gamma}(u,s)(\boldsymbol{\Gamma}(s,s))^{+}\\ &= \frac{s}{u}\boldsymbol{\Gamma}(u,u)(\boldsymbol{\Gamma}(u,u))^{+}\boldsymbol{\Gamma}(u,u)(\boldsymbol{\Gamma}(s,s))^{+}\\ &= \frac{s}{u}\boldsymbol{\Gamma}(u,u)(\boldsymbol{\Gamma}(s,s))^{+} = \boldsymbol{\Gamma}(u,s)(\boldsymbol{\Gamma}(s,s))^{+} = \mathbf{A}(u,s),\end{aligned}$$

where the first equality follows by Eq. (4.2) when $t = u$, the second by Eq. (2.3), the third by Property I of a pseudoinverse from Section 3.5.1 of Petersen and Pedersen [37], the fourth by Eq. (2.3) again, and the last by Eq. (4.2) again.

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

If $0 < s \le t < u \le T$, then Lemma 1 implies that $\boldsymbol{\Gamma}(t,t)$ is positive definite, so that $(\boldsymbol{\Gamma}(t,t))^{+} = (\boldsymbol{\Gamma}(t,t))^{-1}$. Then,

$$\begin{aligned}\mathbf{A}(u,t)\mathbf{A}(t,s) &= \boldsymbol{\Gamma}(u,t)(\boldsymbol{\Gamma}(t,t))^{-1}\boldsymbol{\Gamma}(t,s)(\boldsymbol{\Gamma}(s,s))^{+} \\ &= \frac{s}{t}\boldsymbol{\Gamma}(u,t)(\boldsymbol{\Gamma}(t,t))^{-1}\boldsymbol{\Gamma}(t,t)(\boldsymbol{\Gamma}(s,s)) = \mathbf{A}(u,s),\end{aligned}$$

where Eq. (2.3) is used to obtain the second and third equalities. Since Eq. (4.1) then holds for all $0 < s \le t \le u \le T$, $\mathbf{X}$ is a wide-sense Markov process (as well as a Gaussian process) and hence a (strict-sense) Markov process.

If $\boldsymbol{X}_{\varepsilon}$ is defined as at the start of Section 2, then

$$\boldsymbol{\Gamma}_{\varepsilon}(t,s) \equiv E[\boldsymbol{X}_{\varepsilon}(t)\boldsymbol{X}_{\varepsilon}^{\mathrm{T}}(s)] = \boldsymbol{\Gamma}(t+\varepsilon, s+\varepsilon) - \boldsymbol{\Gamma}(t+\varepsilon,\varepsilon) - \boldsymbol{\Gamma}(\varepsilon, s+\varepsilon) + \boldsymbol{\Gamma}(\varepsilon,\varepsilon) \quad \textbf{(4.16)}$$

for all $0 \le s \le t \le T-\varepsilon$ and any $0 \le \varepsilon < T$ by the definition of a covariance kernel. Substituting Eq. (2.3) into Eq. (4.16), we easily verify that $\boldsymbol{\Gamma}_{\varepsilon}(t,s) = \boldsymbol{\Gamma}(t,s)$. Since the distribution of a Gaussian process is determined by its covariance kernel, this shows that $\boldsymbol{X}$ has stationary increments. ■

Proof: Applying Definition 8.3.3 on page 200 of Adler [1], a Gaussian process $\boldsymbol{X}$ with stationary increments and the covariance kernel for each component as described in Corollary 1 is an index 1/2 Gaussian field. By Theorem 8.3.2 on page 202 Adler [1], it therefore satisfies a uniform stochastic Hölder condition w.p.1, which in turn implies that its sample paths are continuous w.p.1. ■

Proof: The first equality follows immediately from Eq. (2.3). Let $\Gamma_i(t,s) \equiv E[X_i(t)X_i^{\mathrm{T}}(s)]$. The second equality then follows immediately from Corollary 1 since

$$\begin{aligned}&\mathrm{Cov}(X_i(t+s) - X_i(t), X_i(t+u+s) - X_i(t+u)) \\ &\quad = \Gamma_i(s+t, s+t+u) - \Gamma_i(t, s+t+u) - \Gamma_i(s+t, t+u) - \Gamma_i(t, t+u)\end{aligned}$$

and

$$\begin{aligned}&\sqrt{\mathrm{Var}(X_i(t+s) - X_i(t))}\sqrt{\mathrm{Var}(X_i(t+u+s) - X_i(t+u))} \\ &\quad = \Gamma_i(s+t, s+t) - 2\Gamma_i(t, s+t) + \Gamma_i(t,t).\end{aligned}$$ ■

Proof: The SDE in Eq. (2.4) is an example of a linear SDE, the properties of which are well-known and summarized on pages 354–355 of Karatzas and Shreve [31]; see also Theorem 4.10 on pages 144–147 of Lipster and Shiryayev [34] or Section 5 of Beutler [6]. By Eq. (6.6) on page 354 there, its solution $\boldsymbol{X}$ is a Gaussian process. By Eq. (6.11) on page 355 there,

$$\boldsymbol{\Gamma}(t,s) = \boldsymbol{f}(t)\boldsymbol{f}^{-1}(s)\boldsymbol{g}(s) \quad \text{for } 0 \le s \le T, \quad \textbf{(4.17)}$$

where $\boldsymbol{g}(s) \equiv \boldsymbol{\Gamma}(s,s)$. By Eq. (6.3) on page 354 there, $\boldsymbol{f}(\cdot)$ uniquely solves

$$\dot{\boldsymbol{f}}(t) = \boldsymbol{\mu}(t)\boldsymbol{f}(t) \quad \text{for } 0 \le t \le T, \quad \textbf{(4.18)}$$

$$\boldsymbol{f}(0) = \boldsymbol{I}.$$ ■

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

By Eq. (6.13) on page 355 there, $\boldsymbol{g}(\cdot)$ uniquely solves

$$\dot{\boldsymbol{g}}(s) = \boldsymbol{\mu}(s)\boldsymbol{g}(s) + \boldsymbol{g}(s)\boldsymbol{\mu}^{\mathrm{T}}(s) + \boldsymbol{\sigma}(s)\boldsymbol{\sigma}^{\mathrm{T}}(s), \tag{4.19}$$

$$\boldsymbol{g}(0) = \mathbf{0}.$$

When Eqs. (2.5) and (2.6) hold, $\boldsymbol{f}(t) = \boldsymbol{I} - \boldsymbol{\beta}\boldsymbol{\alpha}^{-1}t$ and $\boldsymbol{g}(s) = s(\boldsymbol{\alpha} - \boldsymbol{\beta}s)$ satisfy Eqs. (4.18) and (4.19), and Eq. (2.3) follows from Eq. (4.17).

Proof: By definition, a process is Gaussian if its finite-dimensional distributions are Gaussian. For any set of random variables with a joint Gaussian distribution, it is well-known that the conditional distribution of any subset—given the values of the others—is Gaussian, see page 522 of Rao [39]. Therefore, the process $\{\boldsymbol{B}(t) : 0 \le t \le u \le T\}$ is Gaussian. Its conditional mean is

$$E[\boldsymbol{Z}(t)|\boldsymbol{Z}(u)] = t\boldsymbol{\omega} + \boldsymbol{\Gamma}(t,u)\boldsymbol{\Gamma}(u,u)^{-}(\boldsymbol{Z}(u) - u\boldsymbol{\omega}) = t\boldsymbol{\omega} + tu^{-1}\boldsymbol{\Gamma}(u,u)\boldsymbol{\Gamma}(u,u)^{-}(\boldsymbol{Z}(u) - u\boldsymbol{\omega})$$

for $0 \le t \le u < T$ where the first equality follows from the result from Section 6.2.2 of Puntaten and Styan [38], and the second from Eq. (2.3). Because $E[\boldsymbol{Z}(u)|\boldsymbol{Z}(u)] = \boldsymbol{Z}(u)$ w.p.1, setting $t = u$ in the first equality shows that $\boldsymbol{\Gamma}(u,u)\boldsymbol{\Gamma}(u,u)^{-}(\boldsymbol{Z}(u) - u\boldsymbol{\omega}) = \boldsymbol{Z}(u) - u\boldsymbol{\omega}$ w.p.1. Substituting that result into the final expression on the right-hand side of the second equation yields

$$E[\boldsymbol{Z}(t)|\boldsymbol{Z}(u)] = t\boldsymbol{\omega} + tu^{-1}(\boldsymbol{Z}(u) - u\boldsymbol{\omega}) = tu^{-1}\boldsymbol{Z}(u). \qquad \blacksquare$$

The conditional covariance matrix is

$$\begin{aligned} &E[(\boldsymbol{Z}(t) - E[\boldsymbol{Z}(t)|\boldsymbol{Z}(u)])(\boldsymbol{Z}(s) - E[\boldsymbol{Z}(s)|\boldsymbol{Z}(u)])^{\mathrm{T}}|\boldsymbol{Z}(u)] \\ &\quad = \boldsymbol{\Gamma}(t,s) - \boldsymbol{\Gamma}(t,u)\boldsymbol{\Gamma}(u,u)^{-}\boldsymbol{\Gamma}(u,s) \\ &\quad = \boldsymbol{\Gamma}(t,s) - stu^{-2}\boldsymbol{\Gamma}(u,u)\boldsymbol{\Gamma}(u,u)^{-}\boldsymbol{\Gamma}(u,u) \\ &\quad = \boldsymbol{\Gamma}(t,s) - stu^{-2}\boldsymbol{\Gamma}(u,u) = (\boldsymbol{I} - tu^{-1}\boldsymbol{I})s\boldsymbol{\alpha}, \end{aligned}$$

where the first equality also follows from the results in Section 6.2.2 of Puntaten and Styan [38], the second equality from Eq. (2.3), the third from Definition 2.2 of a generalized inverse in Rao and Mitra [40], and the last by Eq. (2.3) again.

Proof: The random vector of dimension $m(n-1) \times 1$ that is constructed by concatenating $\boldsymbol{B}(t_1), \boldsymbol{B}(t_2), \ldots, \boldsymbol{B}(t_{n-1})$ has a Gaussian distribution. By Lemma 1, its mean is equal to $\mathbf{vec}(\mathbf{M})$, where the $\mathbf{vec}$ operator is defined in Section 10.2 of Petersen and Pedersen [37], and $\mathbf{M}$ is the matrix with jth column equal to $t_j t_n^{-1}\boldsymbol{x}_n$ for $j = 1, \ldots, n-1$. Lemma 1 also implies that the covariance matrix for the random vector is equal to the Kronecker product $\boldsymbol{T} \otimes \boldsymbol{\alpha}$, where the $\otimes$ operator is also defined in Section 10.2 of Petersen and Pedersen [37], and $\boldsymbol{T}$ is the symmetric matrix of dimension $(n-1) \times (n-1)$ with (i,j)th element equal to $t_i(1 - t_j/t_n)$ for $i \le j$. Let $\boldsymbol{Y} \equiv \boldsymbol{X} - \mathbf{M}$, where $\boldsymbol{X}$ is the matrix with jth column equal to $\boldsymbol{x}_j$ for $j = 1, \ldots, n-1$, and $\boldsymbol{y} \equiv \mathbf{vec}(\boldsymbol{Y})$. If

$$p_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-1} | \boldsymbol{x}_n) \equiv \frac{P(\boldsymbol{B}(t_1) \in d\boldsymbol{x}_1, \ldots, \boldsymbol{B}(t_{n-1}) \in d\boldsymbol{x}_{n-1})}{d\boldsymbol{x}_1, \ldots, d\boldsymbol{x}_{n-1}}, \tag{4.20}$$

then its log-likelihood function is

$$\mathcal{L}_1 = \text{const} - \frac{1}{2}\log(|\boldsymbol{T} \otimes \boldsymbol{\alpha}|) - \frac{1}{2}\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{T}^{-1} \otimes \boldsymbol{\alpha}^{-1})\boldsymbol{y}, \tag{4.21}$$

where $|\cdot|$ in this context denotes the matrix determinate operator. The notation on the left-hand side of Eq. (4.20) is suggestive of our later application of this lemma, but for now is simply shorthand for the right-hand side.

Let

$$\boldsymbol{\alpha} = \sum_{j=1}^{m} \boldsymbol{v}_j \boldsymbol{v}_j^{\mathrm{T}} \tag{4.22}$$

denote a decomposition of the real symmetric positive definite matrix $\boldsymbol{\alpha}$ in which each of the $\boldsymbol{v}_i{}'s$ is a column vector. Such a decomposition always exists by the Spectral Theorem. Also, let $\boldsymbol{z} \equiv (\boldsymbol{T}^{-1} \otimes \boldsymbol{\alpha}^{-1})\boldsymbol{y} = (\boldsymbol{T}^{-1} \otimes \boldsymbol{\alpha}^{-1})\mathbf{vec}(\boldsymbol{Y})$. By Eq. (520) of Peterson and Pedersen [37],

$$\boldsymbol{z} = \mathbf{vec}(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}). \tag{4.23}$$

Then,

$$\begin{aligned}
\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{T}^{-1} \otimes \boldsymbol{\alpha}^{-1})\boldsymbol{y} = \boldsymbol{z}^{\mathrm{T}}(\boldsymbol{T} \otimes \boldsymbol{\alpha})\boldsymbol{z} &= \sum_{j=1}^{m} \boldsymbol{z}^{\mathrm{T}}(\boldsymbol{T} \otimes (\boldsymbol{v}_j \boldsymbol{v}_j^{\mathrm{T}}))\boldsymbol{z} \\
&= \sum_{j=1}^{m} \boldsymbol{v}_j^{\mathrm{T}} \boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1}\boldsymbol{v}_j \\
&= tr\left(\sum_{i=j}^{m} \boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1}\boldsymbol{v}_j\boldsymbol{v}_j^{\mathrm{T}}\right) \\
&= tr(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}),
\end{aligned} \tag{4.24}$$

where the first equality follows from the definition of z, the second from Eq. (4.22) here and Eq. (506) of Petersen and Pederson [37], the third from Eq. (4.23) here and Eq. (524) of Petersen and Pedersen [37], the forth from well-known properties of the matrix trace, and the fifth from Eq. (4.22) again.

By Eq. (4.21) and Eq. (4.24),

$$\mathcal{L}_1 = \text{const} - \frac{1}{2}\log(|\boldsymbol{T} \otimes \boldsymbol{\alpha}|) - \frac{1}{2}tr(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}) \tag{4.25}$$

from which we obtain, for each $1 \leq i \leq j \leq m$,

$$\begin{aligned}
\frac{\partial}{\partial \alpha_{i,j}}\mathcal{L}_1 &= -\frac{1}{2}tr\left((\boldsymbol{T}^{-1} \otimes \boldsymbol{\alpha}^{-1})\left(\boldsymbol{T} \otimes \frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right)\right) + \frac{1}{2}tr\left(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1}\frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right) \\
&= -\frac{1}{2}tr\left(\boldsymbol{I}_{(n-1)\times(n-1)} \otimes \boldsymbol{\alpha}^{-1}\frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right) + \frac{1}{2}tr\left(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1}\frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right) \\
&= \frac{1}{2}tr\left(-(n-1)\boldsymbol{\alpha}^{-1}\frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right) + \frac{1}{2}tr\left(\boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1}\frac{\partial \boldsymbol{\alpha}}{\partial \alpha_{i,j}}\right)
\end{aligned} \tag{4.26}$$

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

for $1 \le i \le j \le m$, where $\boldsymbol{I}_{(n-1)\times(n-1)}$ denotes the identity matrix of dimension $(n-1)\times(n-1)$. The first equality of Eq. (4.26) follows by applying Eqs. (57), (512), and (137) of Petersen and Pederson [37] to the first nonconstant summand on the right-hand side of Eq. (4.25) and by applying Eqs. (124) and (137) of Petersen and Pederson [37] to the second nonconstant summand of Eq. (4.25). Both of those steps use the fact that $\boldsymbol{T}$ and $\boldsymbol{\alpha}$ are symmetric. The second equality of Eq. (4.26) follows from Eq. (511) of Petersen and Pederson [37], and the last from the definition of a Kronecker product. ■

Any symmetric matrix $\boldsymbol{\chi}$ of dimension $m \times m$ can be expressed as a linear combination of the $(\partial\boldsymbol{\alpha}/\partial\alpha_{i,j})'s$. Therefore, using Eq. (4.26), $(\partial/\partial\alpha_{i,j})\mathcal{L}_1 = 0$ for all $1 \le i \le j \le m$ if and only if

$$tr((-(n-1)\boldsymbol{\alpha}^{-1} + \boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1})\boldsymbol{\chi}) = 0 \tag{4.27}$$

for any symmetric matrix $\boldsymbol{\chi}$ of dimension $m \times m$ That is true, in particular, for $\boldsymbol{\chi} = (-(n-1)\boldsymbol{\alpha}^{-1} + \boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1})^{\mathrm{T}}$, in which case Eq. (4.27) implies that

$$-(n-1)\boldsymbol{\alpha}^{-1} + \boldsymbol{\alpha}^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\alpha}^{-1} = 0. \tag{4.28}$$

Conversely, Eq. (4.28) implies that Eq. (4.27) holds for any symmetric $\boldsymbol{\chi}$. We can, therefore, conclude that the likelihood equations are uniquely satisfied by

$$\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}} \equiv (n-1)^{-1}\boldsymbol{Y}\boldsymbol{T}^{-1}\boldsymbol{Y}^{\mathrm{T}}. \tag{4.29}$$

We easily verify that $\boldsymbol{T}^{-1}$ is the symmetric tridiagonal matrix

$$\boldsymbol{T}^{-1} = \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & a_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-3} & b_{n-2} & a_{n-2} \\ & & & a_{n-2} & b_{n-1} \end{bmatrix}, \tag{4.30}$$

where $a_j = -1/(t_{j+1}-t_j)$ and $b_j = (t_{j+1}-t_{j-1})/(t_{j+1}-t_j)(t_j-t_{j-1})$ when $t_0 \equiv 0$. For $k = 1, 2, \ldots, m$, let $y_{k,i}$ denote the (k,i)th element $\boldsymbol{Y}$ for $i = 1, 2, \ldots, n-1$, and let $y_{k,n} \equiv 0$. Then, by Eqs. (4.29) and (4.30), the (j,k)th element of $\hat{\boldsymbol{\alpha}}$ is

$$\begin{aligned}
\hat{\alpha}_{j,k} &= \frac{1}{n-1}\sum_{i=1}^{n-1}\left(\frac{t_{i+1}-t_{i-1}}{(t_{i+1}-t_i)(t_i-t_{i-1})}y_{j,i}y_{k,i} - \frac{1}{(t_{i+1}-t_i)}y_{j,i}y_{k,i+1} - \frac{1}{(t_{i+1}-t_i)}y_{j,i+1}y_{k,i}\right) \\
&= \frac{1}{n-1}\left(\frac{t_0 y_{j,1}y_{k,1}}{(t_1^2-t_0t_1)} - \frac{t_{n-1}y_{j,n}y_{k,n}}{(t_n^2-t_{n-1}t_n)} + \sum_{i=1}^{n-1}\frac{(t_{i+1}y_{j,i}-t_iy_{j,i+1})(t_{i+1}y_{k,i}-t_iy_{k,i+1})}{t_it_{i+1}(t_{i+1}-t_i)}\right) \\
&= \frac{1}{n-1}\sum_{i=1}^{n-1}\frac{(t_{i+1}y_{j,i}-t_iy_{j,i+1})(t_{i+1}y_{k,i}-t_iy_{k,i+1})}{t_it_{i+1}(t_{i+1}-t_i)} \\
&= \frac{1}{n-1}\sum_{i=1}^{n-1}\frac{(t_{i+1}x_{j,i}-t_ix_{j,i+1})(t_{i+1}x_{k,i}-t_ix_{k,i+1})}{t_it_{i+1}(t_{i+1}-t_i)}.
\end{aligned}$$

We recognize the expression on the right-hand side of the last equality as the (j,k)th element of Eq. (2.7). Hence, Eqs. (4.29) and (2.7) are equivalent. By Eq. (2.7), $n \ge m+1$ in order for $\hat{\boldsymbol{\alpha}}$ to have rank m.

Applying Eqs. (124), (125), and (137) of Peterson and Pedersen [37] to the expression on the right-hand side of Eq. (4.26) and substituting $\hat{\boldsymbol{\alpha}}$ from Eq. (4.29) results in

$$\frac{\partial^2}{\partial\alpha_{i,j}\partial\alpha_{k,l}}\mathcal{L}_1(\hat{\boldsymbol{\alpha}}) = -\frac{1}{2}(n-1)tr\left(\left(\hat{\boldsymbol{\alpha}}^{-1}\frac{\partial\boldsymbol{\alpha}}{\partial\alpha_{i,j}}\right)\left(\hat{\boldsymbol{\alpha}}^{-1}\frac{\partial\boldsymbol{\alpha}}{\partial\alpha_{k,l}}\right)\right) \quad \textbf{(4.31)}$$

for $1 \le i \le j \le m$ and $1 \le k \le l \le m$. The matrices $\hat{\boldsymbol{\alpha}}^{-1}(\partial\boldsymbol{\alpha}/\partial\alpha_{i,j})$ for $1 \le i \le j \le m$ in Eq. (4.31) are linearly independent, and the matrix trace in Eq. (4.31) is an inner product of two such elements. Consequently, the likelihood function's Hessian matrix evaluated at $\hat{\boldsymbol{\alpha}}$, as expressed by the left-hand side of Eq. (4.31), is the negative of a Gram matrix of independent vectors and is therefore negative definite. That shows that $\mathcal{L}_1(\hat{\boldsymbol{\alpha}})$ is indeed a maximum Eq. (4.21), so that $\hat{\boldsymbol{\alpha}}$ in Eq. (2.7) is the unique maximum-likelihood estimator for $p_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-1}|\boldsymbol{x}_n)$ in Eq. (4.20).

By Eq. (4.25), the log-likelihood function $\mathcal{L}_1$ depends on the sample vectors only through $\hat{\boldsymbol{\alpha}}$ in Eq. (4.29). The sufficiency of $\hat{\boldsymbol{\alpha}}$ then follows from the Fisher–Neyman theorem; see Halmos and Savage [23] for a formulation of that theorem covering the current case of dependent random vectors.

The above proof of Lemma 3 uses Petersen and Pedersen [37] as a common reference for all matrix identities. An alternative proof is outlined by Dutilleul [16] in the context of a related problem. A log-likelihood equation corresponding to Eq. (4.25) here is given at the start of Section 3 there and a solution corresponding to Eq. (4.29) here is given by Eq. (3.4) there. The uniqueness of that solution then follows from Section 18.24 of Kendell and Stuart [32], which also shows that the Hessian matrix corresponding to the left-hand side of Eq. (4.31) here will always be negative semi-definite for a broader class of problems. To show that the solution is indeed a maximum requires verifying that the Hessian matrix is negative definite for the problem at hand. That last step was not covered by Dutilleul [16] but is covered by the proof of Lemma 3 here. The spectral representation for the solution in Eq. (2.7) appears to be a new result.

PROOF: Let

$$p_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-1}|\boldsymbol{x}_n) \equiv \frac{P(\boldsymbol{X}(t_1) \in d\boldsymbol{x}_1, \ldots, \boldsymbol{X}(t_{n-1}) \in d\boldsymbol{x}_{n-1}|\boldsymbol{X}(t_n) = \boldsymbol{x}_n)}{d\boldsymbol{x}_1, \ldots, d\boldsymbol{x}_{n-1}}, \quad \textbf{(4.32)}$$

$$p_2(\boldsymbol{x}_n) \equiv \frac{P(\boldsymbol{X}(t_n) \in d\boldsymbol{x}_n)}{d\boldsymbol{x}_n}, \quad \textbf{(4.33)}$$

and

$$p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) \equiv \frac{P(\boldsymbol{X}(t_1) \in d\boldsymbol{x}_1, \ldots, \boldsymbol{X}(t_n) \in d\boldsymbol{x}_n)}{d\boldsymbol{x}_1, \ldots, d\boldsymbol{x}_n}. \quad \textbf{(4.34)}$$

They are related by

$$p(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = p_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-1}|\boldsymbol{x}_n)p_2(\boldsymbol{x}_n). \quad \textbf{(4.35)}$$

In the setting of Theorem 4, estimators are jointly sufficient if they are jointly sufficient with respect to Eq. (4.34) and are the maximum-likelihood estimators if they maximize Eq. (4.34). ■

By Lemma 2, Eq. (4.32) coincides with the density of the same name in Eq. (4.20). Let $\boldsymbol{\Sigma} \equiv \boldsymbol{\Gamma}(t_n, t_n)$. If $\text{Rank}(\boldsymbol{\Sigma}) = m$, then Eq. (4.33) has the familiar form of a multivariate

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

Gaussian density,

$$p_2(\boldsymbol{x}_n) = (2\pi)^{-m/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left(-\frac{1}{2}tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}})\right). \tag{4.36}$$

The density in Eq. (4.36) does not exist when $\mathrm{Rank}(\boldsymbol{\Sigma}) < m$, but, in that case, there exists a matrix $\boldsymbol{N}$ with dimension $m \times (m - \mathrm{Rank}(\boldsymbol{\Sigma}))$ such that $\mathrm{Rank}(\boldsymbol{N}) = m - \mathrm{Rank}(\boldsymbol{\Sigma})$ and $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{\Sigma} = 0$; see page 527 of Rao [39]. Then, $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{X}(t_n)$ has a variance of zero and is, therefore, equal to its mean w.p.1, that is, $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{X}(t_n) = \boldsymbol{N}^{\mathrm{T}}E\boldsymbol{X}(t_n) = 0$ w.p.1. In this case, Section 8a.4 on pages 527–528 of Rao [39] shows that $\boldsymbol{X}(t_n)$ will have the density,

$$p_2(\boldsymbol{x}_n) = \begin{cases} (2\pi)^{-\mathrm{Rank}(\boldsymbol{\Sigma})/2}|\boldsymbol{\Sigma}|_+^{-1/2}\exp\left(-\frac{1}{2}tr(\boldsymbol{\Sigma}^{-}\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}})\right), & \boldsymbol{N}^{\mathrm{T}}\boldsymbol{x}_n = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{4.37}$$

where $|\boldsymbol{\Sigma}|_+$ denotes the product of the nonzero eigenvalues of $\boldsymbol{\Sigma}$ (and where $\boldsymbol{\Sigma}^-$ denotes any generalized inverse of $\boldsymbol{\Sigma}$ as before). If $\mathrm{Rank}(\boldsymbol{\Sigma}) = m$, then $\boldsymbol{N} = 0$, in which case Eq. (4.37) agrees with Eq. (4.36). Therefore, the density in Eq. (4.35) is equal in the general case to the product of Eq. (4.37) and $p_1(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n-1}|\boldsymbol{x}_n)$ from Eq. (4.20).

When Eqs. (2.7) and (2.8) hold, $\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}} = t_n\hat{\boldsymbol{\alpha}} - t_n^2\hat{\boldsymbol{\beta}}$ in Eq. (4.37). Furthermore, the case in Eq. (4.37) where $\boldsymbol{N}^{\mathrm{T}}\boldsymbol{x}_n = 0$ applies if and only if $0 = \boldsymbol{N}^{\mathrm{T}}\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}}\boldsymbol{N} = \boldsymbol{N}^{\mathrm{T}}(t_n\hat{\boldsymbol{\alpha}} - t_n^2\hat{\boldsymbol{\beta}})\boldsymbol{N}$. Consequently, Eq. (4.37) depends on the sample vector $\boldsymbol{x}_n$ only through $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$. The proof of Lemma 3 shows that Eq. (4.20) depends on the sample vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ only through $\hat{\boldsymbol{\alpha}}$. Since both terms on the right-hand side of Eq. (4.35) depend on the sample vectors only through $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$, the joint sufficiency of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ follows from the Fisher–Neyman theorem for any given $\mathrm{Rank}(\boldsymbol{\Sigma})$.

Section 8a.5, page 532 of Rao [39] implies for the case in which $\mathrm{Rank}(\boldsymbol{\Sigma}) = 1$ that

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}} \tag{4.38}$$

is the maximum-likelihood estimator of $\boldsymbol{\Sigma} \equiv \boldsymbol{\Gamma}(t_n, t_n)$ in Eq. (4.37). Then, Eq. (2.3) implies that any estimators $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ satisfying $\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}} = t_n(\hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}t_n)$ will maximize Eq. (4.37). In particular, Eqs. (2.7) and (2.8) do. By Lemma 3, $\hat{\boldsymbol{\alpha}}$ from Eq. (2.7) is the maximum-likelihood estimator of Eq. (4.32). Since $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ from Eq. (2.7) and Eq. (2.8) maximize the terms on the right-hand side of Eq. (4.35) individually, they are the maximum-likelihood estimators of the parameters of Eq. (4.34).

By Eq. (2.3) and Eq. (2.7),

$$\begin{aligned} E\hat{\boldsymbol{\alpha}} &= \frac{1}{n-1}E\sum_{i=1}^{n-1}\frac{t_{i+1}^2\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - t_it_{i+1}\boldsymbol{x}_i\boldsymbol{x}_{i+1}^{\mathrm{T}} - t_it_{i+1}\boldsymbol{x}_{i+1}\boldsymbol{x}_i^{\mathrm{T}} + t_i^2\boldsymbol{x}_{i+1}\boldsymbol{x}_{i+1}^{\mathrm{T}}}{t_it_{i+1}(t_{i+1} - t_i)} \\ &= \frac{1}{n-1}\sum_{i=1}^{n-1}\frac{t_{i+1}^2\boldsymbol{\Gamma}(t_i, t_i) - t_it_{i+1}\boldsymbol{\Gamma}(t_i, t_{i+1}) - t_it_{i+1}\boldsymbol{\Gamma}(t_{i+1}, t_i) + t_i^2\boldsymbol{\Gamma}(t_{i+1}, t_{i+1})}{t_it_{i+1}(t_{i+1} - t_i)} \\ &= \frac{1}{n-1}\sum_{i=1}^{n-1}\boldsymbol{\alpha} = \boldsymbol{\alpha}, \end{aligned}$$

so that $\hat{\boldsymbol{\alpha}}$ is unbiased. Then, by Eqs. (2.3), (2.7), and (2.8),

$$E(t_n(\hat{\boldsymbol{\alpha}} - \hat{\boldsymbol{\beta}}t_n)) = E\boldsymbol{x}_n\boldsymbol{x}_n^{\mathrm{T}} = \boldsymbol{\Gamma}(t_n, t_n) = t_n(\boldsymbol{\alpha} - \boldsymbol{\beta}t_n),$$

so that $\hat{\boldsymbol{\beta}}$ is unbiased as well.

PROOF: Since $\boldsymbol{Z}$ is a ψ-GMP with drift on $[0,T]$, $\boldsymbol{\Gamma}(t,s) = s(\boldsymbol{\alpha} - \boldsymbol{\beta}t)$ for $0 \le s \le t \le T$. When $0 < s < T$, the covariance matrix $\boldsymbol{\Gamma}(s,s) = \boldsymbol{s}(\boldsymbol{\alpha} - \boldsymbol{\beta}s)$ is nonsingular by Lemma 1. Applying the same logic as in the proof of Lemma 2,

$$\begin{aligned} E[\boldsymbol{Z}^{(s)}(t)|\boldsymbol{Z}(s) = \boldsymbol{z}_s] &= E[\boldsymbol{Z}(s+t) - \boldsymbol{Z}(s)|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s] \\ &= (s+t)\boldsymbol{\omega} + \boldsymbol{\Gamma}(t+s,s)\boldsymbol{\Gamma}(s,s)^{-1}(\boldsymbol{z}_s - s\boldsymbol{\omega}) - \boldsymbol{z}_s \\ &= \boldsymbol{\omega}_s t \end{aligned}$$

for $t \ge s$, where $\boldsymbol{\omega}_s$ is defined as in Eq. (2.9); and

$$\begin{aligned} &E[(\boldsymbol{Z}^{(s)}(u) - E[\boldsymbol{Z}^{(s)}(u)|\boldsymbol{Z}(s) = \boldsymbol{z}_s])(\boldsymbol{Z}^{(s)}(t) - E[\boldsymbol{Z}^{(s)}(t)|\boldsymbol{Z}(s) = \boldsymbol{z}_s])|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s] \\ &= \boldsymbol{\Gamma}(u+s,t+s) - \boldsymbol{\Gamma}(u+s,s)\boldsymbol{\Gamma}(s,s)^{-1}\boldsymbol{\Gamma}(s,t+s) \\ &= t(\boldsymbol{\alpha} - \boldsymbol{\beta}_s u) \end{aligned}$$

for $u \ge t \ge s$, where $\boldsymbol{\beta}_s$ is also defined in Eq. (2.9). The result in Lemma 4 then follows because $\boldsymbol{Z}$ is a Markov process. ∎

PROOF: Theorem 5 is the special case when $m = 1$ of Theorem 6. Under the assumptions of Theorem 6, $\{\boldsymbol{Z}(t) \equiv t\boldsymbol{\omega} + \boldsymbol{X}(t) : 0 \le t \le T\}$ where $\boldsymbol{\omega}$ is a constant vector and $\boldsymbol{X}$ is an $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ ψ-GMP on $[0,T]$. ∎

Fixing s, let $\{\boldsymbol{Z}^{(s)}(t) \equiv \boldsymbol{Z}(s+t) - \boldsymbol{Z}(s) : 0 \le t \le T - s\}$. Also, let $Z \equiv Z_i$ denotes the ith component of $\boldsymbol{Z}$, and let $Z^{(s)} \equiv Z_i^{(s)}$ denotes the ith component of $\boldsymbol{Z}^{(s)}$. Fixing u, let

$$H(z_u^{(s)}|\boldsymbol{z}_s, q_s) \equiv P(Z^{(s)}(u) \le z_u^{(s)}|\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s), \quad \textbf{(4.39)}$$

$$F(q_{s+u}|\boldsymbol{z}_s, q_s) \equiv P(Q(s+u) \le q_{s+u}|\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s), \quad \textbf{(4.40)}$$

and

$$F(q_{s+u}|\boldsymbol{z}_s, q_s, z_u^{(s)}) \equiv P(Q(s+u) \le q_{s+u}|\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s, Z^{(s)}(u) = z_u^{(s)}). \quad \textbf{(4.41)}$$

The process $\{Z^{(s)}(t) : 0 \le t \le T - s|\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s\}$ that is obtained by conditioning $Z^{(s)}$ on both $\boldsymbol{Z}(s) = \boldsymbol{z}_s$ and $Q(s) = q_s$ has the same distribution as the process $\{Z^{(s)}(t) : 0 \le t \le T - s|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s\}$ that is obtained by conditioning $Z^{(s)}$ on $\boldsymbol{Z}(s) = \boldsymbol{z}_s$ alone because $\boldsymbol{Z}$ is a Markov process and $Q(s)$ is determined by $\{\boldsymbol{Z}(v) : 0 \le v \le s\}$ by Eq. (2.10). By Lemma 4,

$$E[\boldsymbol{Z}^{(s)}(t)|\boldsymbol{Z}(s) = \boldsymbol{z}_s] = \boldsymbol{\omega}_s t$$

for $t \ge s$, where $\boldsymbol{\omega}_s$ is defined as in Eq. (2.9); and

$$\begin{aligned} &E[(\boldsymbol{Z}^{(s)}(u) - E[\boldsymbol{Z}^{(s)}(u)|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s])(\boldsymbol{Z}^{(s)}(t) - E[\boldsymbol{Z}^{(s)}(t)|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s])|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s] \\ &= t(\boldsymbol{\alpha} - \boldsymbol{\beta}_s u) \end{aligned}$$

for $u \ge t \ge s$, where $\boldsymbol{\beta}_s$ is also defined in Eq. (2.9). By Corollary 1, the $\mathbb{R}^1$- valued process

$$\{Z^{(s)}(t) : 0 \le t \le T - s|\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s\} = \{Z^{(s)}(t) : 0 \le t \le T - s|\boldsymbol{Z}(s) = \boldsymbol{z}_s\}$$

is then an (α, β_s) ψ-GMP with drift ω_s on $[0, T - s]$ where α, ω_s and β_s are defined in Eq. (2.18), so that

$$H(z_u^{(s)}|\boldsymbol{z}_s, q_s) = N(z_u^{(s)}; \omega_s u, u(\alpha - \beta_s u)), \quad \textbf{(4.42)}$$

where $N(y; \mu, \sigma^2) \equiv \int_{-\infty}^{y} \exp(-(x-\mu)^2/(2\sigma^2))/\sqrt{2\pi\sigma^2}dx$ is the cumulative distribution function for a normal random variable with mean μ and variance σ^2.

By the above results and Lemma 2, the process $\{Z^{(s)}(t) : 0 \le t \le u \le T - s|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s,\ Z^{(s)}(u) = z_u^{(s)}\}$ that is obtained by conditioning $\{Z^{(s)}(t) : 0 \le t \le u \le T - s|\ \boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s\}$ on $Z^{(s)}(u) = z_u^{(s)}$ is an $(\alpha, u^{-1}\alpha)$ ψ-GMP with drift $u^{-1}z_u^{(s)}$ on $[0, u]$. Since that process does not depend on ω_s from (2.18), it does not depend on $\boldsymbol{z}_s$.

To summarize, conditioning $Z^{(s)}$ on both $\boldsymbol{Z}(s) = \boldsymbol{z}_s$ and $Q(s) = q_s$ results in an (α, β_s) ψ-GMP with drift ω_s on $[0, T - s]$. That process depends on $\boldsymbol{z}_s$ but not q_s. Further conditioning it on $Z^{(s)}(u) = z_u^{(s)}$ for some $u \le T - s$ results in an $(\alpha, u^{-1}\alpha)$ ψ-GMP with drift $u^{-1}z_u^{(s)}$ on $[0, u]$. That process depends on $z_u^{(s)}$ but on neither q_s nor $\boldsymbol{z}_s$. Therefore, the process obtained by conditioning $\{Q(t) :\ s \le t \le s + u\}$ on $\boldsymbol{Z}(s) = \boldsymbol{z}_s, Q(s) = q_s$, and $Z^{(s)}(u) = z_u^{(s)}$ begins in state q_s at time s and evolves from there according to a net input process that is an $(\alpha, u^{-1}\alpha)$ ψ-GMP with drift $u^{-1}z_u^{(s)}$ on $[0, u]$. That process is a scaled BB. Section 3 of Hajek [22] derived the distribution for the length of a queue with such a net input process. Following the development there, we find that

$$F(q_{s+u} \mid \boldsymbol{z}_s,\ q_s, z_u^{(s)}) = \begin{cases} 1 - e^{-2q_{s+u}(q_{s+u} - z_u^{(s)})/(\alpha \mathrm{u})}, & q_{s+u} \ge \max(q_s + z_u^{(s)}, 0) \\ 0, & \text{otherwise.} \end{cases} \quad \textbf{(4.43)}$$

Then,

$$F(q_{s+u}|\boldsymbol{z}_s,q_s) = \int_{-\infty}^{q_{s+u}-q_s} F(q_{s+u}|\boldsymbol{z}_s,q_s, z_u^{(s)})dH(z_u^{(s)}|\boldsymbol{z}_s,q_s) \quad \text{for } q_{s+u} \ge 0, \quad \textbf{(4.44)}$$

where the limits of integration follow from Eq. (4.43). Using the technique for completing the square in the exponent (demonstrated in the example on page 13 of Harrison [24]), we evaluate Eq. (4.44) using Eqs. (4.42) and (4.43) to obtain Eq. (2.11).

Proof: The identities in Eqs. (2.13) and (2.14) follow from well-known properties of conditional densities; see, for example, Theorem 8.3.3 on page 142 of Whittle [48]. The expression in Eq. (2.15) describes the density of Eq. (2.11). By the same logic as led to Eq. (4.43) above,

$$P(Q(s) \le q_s|Z(s) = z_s) = \begin{cases} 1 - e^{-2q_s(q_s - z_s)/(\alpha s)}, & q_s \ge \max(Q(0) + z_s, 0) \\ 0, & \text{otherwise.} \end{cases}$$

■

The expression in Eq. (2.16) describes its density. Because Z is an (α, β) ψ-GMP with drift ω, $P(Z(s) \le z_s) = N(z_s; \omega s, s(\alpha - \beta s))$ and the expression in Eq. (2.17) describes its density.

References

1. Adler, R.J. (1981). *The geometry of random fields*. Chichester: Wiley.
2. Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: a close-form approximation approach. *Econometria* 70(1): 223–262.
3. Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Annals of Statistics* 36(2): 906–937.
4. Arthur, W.B., Ermoliev, Y.M., & Kanjovski, Y.M. (1987). Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research* 30: 294–303.
5. Ball, C.A. & Torous, W.N. (1983). Bond pricing dynamics and options. *The Journal of Financial and Quantitative Analysis* 18(4): 517–531.
6. Beutler, F. (1963). Multivariate wide-sense Markov processes and prediction theory. *Annals of Mathematical Statistics* 34: 424–438.

https://doi.org/10.1017/S0269964820000388 Published online by Cambridge University Press

7. Cont, R. (2010). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* 1: 223–236.
8. Dai, J.G. (1992). Reflected Brownian Motion in an orthant: numerical methods for steady-state analysis. *Annals of Probability* 2(1): 65–86.
9. Debicki, K. & Mandjes, M. (2011). Open problems in Gaussian fluid queueing theory. *Queueing Systems* 68: 267–274.
10. Debicki, K. & Rolski, T. (1995). A Gaussian fluid model. *Queuing Systems* 20(3–4): 443–452.
11. Debicki, K. & Rolski, T. (2002). A note on transient Gaussian fluid models. *Queuing Systems* 41(4): 321–342.
12. Debicki, K., Es-Saghouani, A., & Mandjes, M. (2009). Transient characteristics of Gaussian queues. *Queuing Systems* 62: 383–409.
13. Debicki, K., Kosinski, K., & Mandjes, M. (2012). Gaussian queues in light and heavy traffic. *Queuing Systems* 71: 137–149.
14. Doob, J.L. (1944). The elementary Gaussian processes. *Annals of Mathematical Statistics* 15(3): 229–282.
15. Doob, J.L. (1953). *Stochastic processes*. New York: Wiley.
16. Dutilleul, P. (1999). The MLE algorithm for the matrix normal distribution. *Journal of Statistical Computation and Simulation* 64(2): 105–123.
17. Feller, W. (1971). *An introduction to probability theory and its applications*, vol. II, 2nd ed. New York: Wiley.
18. Fendick, K.W. & Whitt, W. (1989). Measurements and approximations to describe the offered traffic and predict the average workload in a single-server queue. *Proceedings of the IEEE* 77: 171–194.
19. Fendick, K.W., Saksena, V.R., & Whitt, W. (1989). Dependence in packet queues. *IEEE Transactions in Communications* 37: 1173–1183.
20. Fendick, K.W., Saksena, V.R., & Whitt, W. (1991). Investigating dependence in packet queues with the index of dispersion for work. *IEEE Transactions in Communications* 39: 1231–1243.
21. Fischer, J.W., Walter, W.D., & Avery, M.L. (2013). Brownian bridge movement models to characterize birds' home ranges. *The Condor* 115(2): 298–305.
22. Hajek, B. (1994). A queue with periodic arrivals and constant service rate. In F.P. Kelly (ed.), *Probability, statistics, and optimisation: a tribute to Peter Whittle*. Chichester: Wiley, pp. 147–157.
23. Halmos, P.R. & Savage, L.J. (1949). Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *Annals of Mathematical Statistics* 20(2): 225–241.
24. Harrison, J.M. (1985). *Brownian motion and stochastic flow systems*. New York: Wiley.
25. Harrison, J.M. & Reiman, M.I. (1981). Reflected Brownian motion on an orthant. *Annals of Probability* 9(2): 302–308.
26. Hida, T. (1960). Canonical representations of Gaussian processes and their applications. *Memoirs of the College of Science, University of Kyoto, A* 33: 109–155.
27. Higham, N.J. (1988). Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra and its Applications* 103: 103–118.
28. Higham, N.J. (2002). Computing the nearest correlation matrix – a problem in finance. *IMA Journal of Numerical Analysis* 22(3): 329–343.
29. Horne, J.S. (2007). Analyzing animal movements using Brownian bridges. *Ecology* 88(9): 2354–2363.
30. Iglehart, D.L. & Whitt, W. (1970). Multiple channel queues in heavy traffic, II: sequences, networks, and batches. *Advances in Applied Probability* 2: 355–369.
31. Karatzas, I. & Shreve, S. (2000). *Brownian motion and stochastic calculus*, 2nd ed. New York: Springer.
32. Kendall, M.G. & Stuart, A. (1983). *The advanced theory of statistics*, vol. 2. Belmont: Wadsworth.
33. Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10: 603–621.
34. Lipster, R.S. & Shiryayev, A.N. (1974). *Statistics of random processes I: general theory*. New York: Springer-Verlag.
35. Mandelbrot, B. & van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10(4): 422–437.
36. Mandrekar, V. (1968). On multivariate wide-sense Markov processes. *Nagoya Mathematical Journal* 33: 7–19.
37. Petersen, K.B. & Pedersen, M.S. (2012). *The matrix cookbook*. Kongens Lyngby: Technical University of Denmark.
38. Puntanen, S. & Styan, G.P. (2005). Schur complements in statistics and probability. In Zhang, F. (ed.), *Schur complement and its applications, numerical methods and algorithms*, vol. 4. Boston: Springer, pp. 163–226.

39. Rao, C. (1973). *Linear statistical inference and its applications*. New York: Wiley.
40. Rao, C. R. & Mitra, S. K. (1972). Generalized inverse of a matrix and is applications. In L. Le Cam, J. Neyman, E. Scott (eds.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. Berkeley: University of California Press.
41. Reiman, M.I. (1984). Open queuing networks in heavy traffic. *Mathematics of Operations Research* 9: 441–458.
42. Schäfer, J. & Strimmer, K. (2005). A shrinkage approach to large scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology* 4(1), Article 32.
43. Sriram, K. & Whitt, W. (1986). Characterizing superposition arrival processes in packet multiplexers for voice and data. *IEEE Journal on Selected Areas of Communication* 4(6): 833–846.
44. Whitt, W. (1983). The queuing network analyzer. *The Bell System Technical Journal* 62(9): 2779–2815.
45. Whitt, W. (2002). *Stochastic-process limits*. New York: Springer.
46. Whitt, W. & You, W. (2018). Using robust queuing to expose the impact of dependence in single-server queues. *Operations Research* 66(1): 184–199.
47. Whitt, W. & You, W. (2019). *Time-varying Robust Queueing*. (Operations Research) https://pubsonline.informs.org/doi/10.1287/opre.2019.1846.
48. Whittle, P. (1992). *Probability via expectation*, 3rd ed. New York: Springer-Verlag.
49. Yaglom, A.M. (1962). *Theory of stationary random functions*. Englewood Cliffs: Prentice-Hall.