THE DISTRIBUTION OF SAMPLE MEAN-VARIANCE PORTFOLIO WEIGHTS

Raymond Kan, Nathan Lassance, Xiaolu Wang

LIDAM Discussion Paper LFIN 2023 / 06

UCLouvain

LiDAM
LFIN | Louvain Institute of Data Analysis and Modeling in economics and statistics

LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html

THE DISTRIBUTION OF SAMPLE MEAN-VARIANCE PORTFOLIO WEIGHTS

RAYMOND KAN

Rotman School of Management, University of Toronto
raymond.kan@rotman.utoronto.ca

NATHAN LASSANCE

LFIN/LIDAM, UCLouvain
nathan.lassance@uclouvain.be

XIAOLU WANG

Ivy College of Business, Iowa State University
xiaoluw@iastate.edu

We present a simple stochastic representation for the joint distribution of sample estimates of three scalar parameters and two vectors of portfolio weights that characterize the minimum-variance frontier. This stochastic representation is useful for sampling observations efficiently, deriving moments in closed-form, and studying the distribution and performance of many portfolio strategies that are functions of these five variables. We also present the asymptotic joint distributions of these five variables for both the standard regime and the high-dimensional regime. Both asymptotic distributions are simpler than the finite-sample one, and the one for the high-dimensional regime, i.e., when the number of assets and the sample size go together to infinity at a constant rate, reveals the high-dimensional properties of the considered estimators. Our results extend upon [T. Bodnar, H. Dette, N. Parolya and E. Thorstén, Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions, *Random Matrices: Theory Appl.* **11** (2022)].

Keywords: portfolio choice; estimation risk; stochastic representation; high-dimensional asymptotics; minimum-variance frontier.

Mathematics Subject Classification 2000: 62E15, 62E20, 62H10, 62H12, 62P05, 91B28.

1. Introduction

Consider an investment universe of $N \geq 2$ assets and a sample of historical excess returns of size T, $(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T)$, where each vector $\boldsymbol{r}_t$ is of dimension $N \times 1$ and $\boldsymbol{r}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.[a] We assume the covariance matrix $\boldsymbol{\Sigma}$ is positive definite and thus invertible. We define the weights of the global minimum-variance portfolio and a zero-cost portfolio as

$$\boldsymbol{w}_g := \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}_N}{\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_N}, \tag{1.1}$$

$$\boldsymbol{w}_z := \boldsymbol{Q}\boldsymbol{\mu}, \tag{1.2}$$

where $\mathbf{1}_N$ is an $N \times 1$ vector of ones and

$$\boldsymbol{Q} := \boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}_N\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}}{\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_N}. \tag{1.3}$$

Note that $\boldsymbol{w}_g^{\mathrm{T}}\mathbf{1}_N = 1$ and $\boldsymbol{w}_z^{\mathrm{T}}\mathbf{1}_N = 0$. Any linear combination of $\boldsymbol{w}_g$ and $\boldsymbol{w}_z$ of the form $\boldsymbol{w}_g + c\boldsymbol{w}_z$ with $c \in \mathbb{R}$ delivers a portfolio on the minimum-variance frontier. This minimum-variance frontier can be completely characterized by the following three parameters:

$$\sigma_g^2 := \frac{1}{\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_N}, \tag{1.4}$$

$$\mu_g := \frac{\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}_N^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{1}_N}, \tag{1.5}$$

$$\psi^2 := \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{\mu}. \tag{1.6}$$

The parameters (μ_g, σ_g^2) are the mean and the variance of the return of $\boldsymbol{w}_g$, and $\psi \geq 0$ is the slope of the asymptote to the minimum-variance frontier.

The quantities $(\sigma_g^2, \mu_g, \psi^2, \boldsymbol{w}_g, \boldsymbol{w}_z)$ together allow us to study efficient mean-variance portfolios and their performance. However, they are unobservable in practice because $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown. To estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, we can use their sample counterparts:

$$\hat{\boldsymbol{\mu}} := \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}_t, \tag{1.7}$$

$$\hat{\boldsymbol{\Sigma}} := \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})^{\mathrm{T}}. \tag{1.8}$$

We assume throughout that $T > N$, and thus $\hat{\boldsymbol{\Sigma}}$ is invertible. Under the i.i.d. multivariate normality assumption on $\boldsymbol{r}_t$, we have that $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$, $T\hat{\boldsymbol{\Sigma}} \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, and they are mutually independent.

[a]Our use of excess returns instead of raw returns follows [1] and is without loss of generality because one can set the risk-free rate equal to zero and work with raw returns instead.

Using $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, we obtain the following sample estimates of $(\sigma_g^2, \mu_g, \psi^2, \boldsymbol{w}_g, \boldsymbol{w}_z)$:

$$\hat{\sigma}_g^2 := \frac{1}{\mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N}, \tag{1.9}$$

$$\hat{\mu}_g := \frac{\mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}}{\mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N}, \tag{1.10}$$

$$\hat{\psi}^2 := \hat{\boldsymbol{\mu}}^{\mathrm{T}} \hat{\boldsymbol{Q}} \hat{\boldsymbol{\mu}}, \tag{1.11}$$

$$\hat{\boldsymbol{w}}_g := \frac{\hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N}{\mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N}, \tag{1.12}$$

$$\hat{\boldsymbol{w}}_z := \hat{\boldsymbol{Q}} \hat{\boldsymbol{\mu}}, \tag{1.13}$$

where $\hat{\boldsymbol{Q}} := \hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} / (\mathbf{1}_N^{\mathrm{T}} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}_N)$.

Our objective in this paper is threefold. First, we present a simple finite-sample stochastic representation of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ that involves $3N - 1$ random variables. The weights of many portfolio strategies proposed in the literature are functions of these five variables and, with our stochastic representation, we can study any function of these portfolio weights, e.g., the linear function such as the out-of-sample portfolio mean return and the quadratic function such as the out-of-sample portfolio return variance.

Second, we show that the number of random variables can be reduced significantly when one needs a stochastic representation of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$, where $\boldsymbol{L}$ is a constant matrix of size $k \times N$ for any positive integer k. Specifically, we present a stochastic representation that involves only $3m + 5$ random variables where $m := \mathrm{rank}(\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}^{\mathrm{T}})$, and use this representation to derive the mean and covariance matrix of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$. Our representation extends upon [2] by allowing a more general $\boldsymbol{L}$ matrix: [2] require $k \leq N - 2$ and $m = k$ (i.e., $\boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}^{\mathrm{T}}$ is of full rank), whereas we do not constrain k and can have $m < k$.[b]

Third, we present the asymptotic distribution of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ in two different regimes: the standard regime in which N is fixed while $T \to \infty$, and the high-dimensional regime in which $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in (0, 1)$. The latter reveals the high-dimensional properties of the five random variables and can serve as a simple approximation to the more involved finite-sample joint distribution.

2. Stochastic representations

In Sections 2.1 and 2.2, we present our stochastic representations for $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ and $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$, respectively. The proofs of all results in the paper are available in Appendix A.

[b]Note that the restriction in [2] means that they do not cover, for example, the case $\boldsymbol{L} = \boldsymbol{I}_N$ with $\boldsymbol{I}_N$ being the $N \times N$ identity matrix.

2.1. *Stochastic representation of* $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$

In the next theorem, we present a stochastic representation for $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ that involves $3N-1$ random variables. We denote by $t_\nu(\boldsymbol{S})$ a multivariate central t-distribution with ν degrees of freedom and a scaling matrix $\boldsymbol{S}$.

Theorem 2.1. *Let $\boldsymbol{P}$ be an $N \times (N-1)$ orthonormal matrix whose columns are orthogonal to $\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{1}_N$. Let $a \sim \mathcal{N}(0,1)$, $\boldsymbol{y} \sim \mathcal{N}(\sqrt{T}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}, \boldsymbol{I}_{N-1})$, $z \sim \mathcal{N}(\sqrt{T}\mu_g/\sigma_g, 1)$, $v_1 \sim \chi^2_{T-N}$, $v_2 \sim \chi^2_{T-N+1}$, $\boldsymbol{t}_1 \sim t_{T-N+2}(\boldsymbol{I}_{N-2})/\sqrt{T-N+2}$, $\boldsymbol{t}_2 \sim t_{T-N+3}(\boldsymbol{I}_{N-2})/\sqrt{T-N+3}$, and they are mutually independent. Let $u := \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} \sim \chi^2_{N-1}(T\psi^2)$ and $\boldsymbol{R}$ be an $(N-1)\times(N-2)$ orthonormal matrix whose columns are orthogonal to $\boldsymbol{y}$.*[c] *Then, $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ are jointly distributed as*

$$\hat{\sigma}_g^2 \stackrel{d}{=} \frac{\sigma_g^2 v_1}{T}, \tag{2.1}$$

$$\hat{\mu}_g \stackrel{d}{=} \frac{\sigma_g}{\sqrt{T}}\left(z + \frac{a\sqrt{u}}{\sqrt{v_2}}\right), \tag{2.2}$$

$$\hat{\psi}^2 \stackrel{d}{=} \frac{u}{v_2}, \tag{2.3}$$

$$\hat{\boldsymbol{w}}_g \stackrel{d}{=} \boldsymbol{w}_g + \sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\left[\frac{a}{\sqrt{v_2 u}}\boldsymbol{y} + \boldsymbol{R}\left(\frac{a}{\sqrt{v_2}}\boldsymbol{t}_1 + (\boldsymbol{I}_{N-2} + \boldsymbol{t}_1\boldsymbol{t}_1^{\mathrm{T}})^{\frac{1}{2}}\boldsymbol{t}_2\right)\right], \tag{2.4}$$

$$\hat{\boldsymbol{w}}_z \stackrel{d}{=} \frac{\sqrt{T}}{v_2}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\left(\boldsymbol{y} + \sqrt{u}\boldsymbol{R}\boldsymbol{t}_1\right). \tag{2.5}$$

Several comments are in order. First, this stochastic representation only involves $3N-1$ random variables, and thus it provides a much more efficient way to sample from the distribution of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ than to directly sample from $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$.

Second, in (2.4) we can write

$$(\boldsymbol{I}_{N-2} + \boldsymbol{t}_1\boldsymbol{t}_1^{\mathrm{T}})^{\frac{1}{2}} = \boldsymbol{I}_{N-2} + \frac{\sqrt{1+\boldsymbol{t}_1^{\mathrm{T}}\boldsymbol{t}_1} - 1}{\boldsymbol{t}_1^{\mathrm{T}}\boldsymbol{t}_1}\boldsymbol{t}_1\boldsymbol{t}_1^{\mathrm{T}}, \tag{2.6}$$

which allows us to use vectorized code for simulating $\hat{\boldsymbol{w}}_g$.

Third, $\hat{\sigma}_g^2$ only depends on v_1, and thus it is independent of $(\hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ that do not depend on v_1.

Fourth, the stochastic representation of $(\hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ simplifies when $N=2$. Indeed, $\boldsymbol{R}$ is a null matrix, $\boldsymbol{y} \equiv y \sim \mathcal{N}(\sqrt{T}\psi, 1)$, $u = y^2$, and thus

$$\hat{\boldsymbol{w}}_g \stackrel{d}{=} \boldsymbol{w}_g + \sigma_g\frac{\mathrm{sgn}(y)a}{\sqrt{v_2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}, \tag{2.7}$$

$$\hat{\boldsymbol{w}}_z \stackrel{d}{=} \frac{\sqrt{T}y}{v_2}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}. \tag{2.8}$$

[c] $\boldsymbol{P}$ and $\boldsymbol{R}$ satisfy $\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{Q}$ and $\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{Q} - \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}/u$.

Finally, using Theorem 2.1, we can obtain stochastic representations for any function of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ which include the weights of many popular portfolios. For example, the sample tangency portfolio, $\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}/(\mathbf{1}_N^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}})$, can be expressed as $\hat{\boldsymbol{w}}_g + (\hat{\sigma}_g^2/\hat{\mu}_g)\hat{\boldsymbol{w}}_z$. The two-fund and three-fund rules studied by [4], $c\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$ and $c\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}} + d\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}_N$, can be written as $c(\hat{\mu}_g/\hat{\sigma}_g^2)\hat{\boldsymbol{w}}_g + c\hat{\boldsymbol{w}}_z$ and $(d + c\hat{\mu}_g)/\hat{\sigma}_g^2\hat{\boldsymbol{w}}_g + c\hat{\boldsymbol{w}}_z$, respectively. The fully invested two-fund rule studied by [1] and [5] is $\hat{\boldsymbol{w}}_g + c\hat{\boldsymbol{w}}_z$. Moreover, the estimates of the optimal combining coefficients (c, d) proposed by [4], [1], and [5] are functions of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2)$, and thus the distribution of the weights of the estimated optimal portfolios can be studied using Theorem 2.1 as well.

Given the stochastic representations of the weights of these sample portfolios $\hat{\boldsymbol{w}}$, we can also study the distribution of any function of $\hat{\boldsymbol{w}}$, e.g., a linear function such as the out-of-sample portfolio mean return $\hat{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\mu}$, and a quadratic function such as the out-of-sample portfolio return variance, $\hat{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}$.

2.2. *Stochastic representation of* $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$

Now, suppose we are interested in the distribution of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$, where $\boldsymbol{L}$ is a $k \times N$ constant matrix and k can be any positive integer. Using the results in Theorem 2.1, we can obtain a stochastic representation that involves $3N-1$ random variables. Instead, we now present an approach that requires a significantly smaller number of random variables. For that purpose, define

$$\boldsymbol{A} := \boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}^{\mathrm{T}} \quad \text{and} \quad m := \operatorname{rank}(\boldsymbol{A}). \tag{2.9}$$

Because $\operatorname{rank}(\boldsymbol{Q}) = N - 1$, we have $m \le N - 1$ even if $k > N - 1$, in which case $\boldsymbol{A}$ is not invertible. We denote the eigen-decomposition of $\boldsymbol{A}$ by

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{\mathrm{T}}, \tag{2.10}$$

where $\boldsymbol{D}$ is the diagonal matrix of the m nonzero eigenvalues and $\boldsymbol{V}$ the $k \times m$ matrix of the corresponding eigenvectors. In the next theorem, we present a stochastic representation for $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ that involves $3m + 5$ random variables.[d]

Theorem 2.2. *Let* $a \sim \mathcal{N}(0,1)$, $\tilde{\boldsymbol{y}} \sim \mathcal{N}(\sqrt{T}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z, \boldsymbol{I}_m)$, $z \sim \mathcal{N}(\sqrt{T}\mu_g/\sigma_g, 1)$, $v_1 \sim \chi^2_{T-N}$, $v_2 \sim \chi^2_{T-N+1}$, $u_0 \sim \chi^2_{N-m-1}(T\psi^2 - T\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z)$ *if* $m < N - 1$ *and* $u_0 = 0$ *if* $m = N - 1$, $\tilde{\boldsymbol{t}}_1 \sim t_{T-N+2}(\boldsymbol{I}_m)/\sqrt{T-N+2}$, $\tilde{\boldsymbol{t}}_2 \sim t_{T-N+3}(\boldsymbol{I}_m)/\sqrt{T-N+3}$, *and they are mutually independent. Let* $u := u_0 + \tilde{\boldsymbol{y}}^{\mathrm{T}}\tilde{\boldsymbol{y}} \sim \chi^2_{N-1}(T\psi^2)$ *and*

$$\boldsymbol{B} := \boldsymbol{I}_m - \frac{1 - \sqrt{\frac{u_0}{u}}}{\tilde{\boldsymbol{y}}^{\mathrm{T}}\tilde{\boldsymbol{y}}}\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{T}}. \tag{2.11}$$

[d]When $m = N - 1$, the random variable $u_0 = 0$ and only $3N + 1$ random variables are needed.

Then, $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ *are jointly distributed as in* (2.1)–(2.3) *and*

$$\boldsymbol{L}\hat{\boldsymbol{w}}_g \overset{d}{=} \boldsymbol{L}\boldsymbol{w}_g + \sigma_g \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \left[\frac{a}{\sqrt{v_2 u}} \tilde{\boldsymbol{y}} + \boldsymbol{B} \left(\frac{a}{\sqrt{v_2}} \tilde{\boldsymbol{t}}_1 + (\boldsymbol{I}_m + \tilde{\boldsymbol{t}}_1 \tilde{\boldsymbol{t}}_1^{\mathrm{T}})^{\frac{1}{2}} \tilde{\boldsymbol{t}}_2 \right) \right], \tag{2.12}$$

$$\boldsymbol{L}\hat{\boldsymbol{w}}_z \overset{d}{=} \frac{\sqrt{T}}{v_2} \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \left(\tilde{\boldsymbol{y}} + \sqrt{u} \boldsymbol{B} \tilde{\boldsymbol{t}}_1 \right). \tag{2.13}$$

Theorem 2.2 is related to [2] and [3], who derive a stochastic representation of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{\eta}})$, where

$$\hat{\boldsymbol{\eta}} := \frac{\hat{\boldsymbol{w}}_z}{\hat{\psi}^2}. \tag{2.14}$$

We extend their results by allowing a more general $\boldsymbol{L}$ matrix. Specifically, they restrict $k \leq N-2$ and $m = k$ (i.e., $\boldsymbol{A}$ is of full rank), while we allow $\boldsymbol{L}$ to be of any dimension and $\boldsymbol{A}$ to have less than full rank.

We can use Theorem 2.2 to derive useful statistics of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ such as the moments. For example, we can show that

$$\mathbb{E}\left[\hat{\sigma}_g^2,\, \hat{\mu}_g,\, \hat{\psi}^2,\, \boldsymbol{L}\hat{\boldsymbol{w}}_g,\, \boldsymbol{L}\hat{\boldsymbol{w}}_z \right] = \left[\sigma_g^2 \tfrac{T-N}{T},\, \mu_g,\, \tfrac{T\psi^2+N-1}{T-N-1},\, \boldsymbol{L}\boldsymbol{w}_g,\, \tfrac{T}{T-N-1} \boldsymbol{L}\boldsymbol{w}_z \right], \tag{2.15}$$

where the expectations of $\hat{\psi}^2$ and $\boldsymbol{L}\hat{\boldsymbol{w}}_z$ exist if $T > N+1$, and

$$\begin{aligned} &\mathbb{V}\mathrm{ar}\left[\hat{\sigma}_g^2,\, \hat{\mu}_g,\, \hat{\psi}^2,\, \boldsymbol{L}\hat{\boldsymbol{w}}_g,\, \boldsymbol{L}\hat{\boldsymbol{w}}_z \right] \\ &= \begin{bmatrix} \mathbb{V}\mathrm{ar}[\hat{\sigma}_g^2] & 0 & 0 & \boldsymbol{0}_k^{\mathrm{T}} & \boldsymbol{0}_k^{\mathrm{T}} \\ 0 & \mathbb{V}\mathrm{ar}[\hat{\mu}_g] & 0 & \mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_g, \hat{\mu}_g]^{\mathrm{T}} & \boldsymbol{0}_k^{\mathrm{T}} \\ 0 & 0 & \mathbb{V}\mathrm{ar}[\hat{\psi}^2] & \boldsymbol{0}_k^{\mathrm{T}} & \mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_z, \hat{\psi}^2]^{\mathrm{T}} \\ \boldsymbol{0}_k & \mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_g, \hat{\mu}_g] & \boldsymbol{0}_k & \mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{w}}_g] & \boldsymbol{0}_{k\times k} \\ \boldsymbol{0}_k & \boldsymbol{0}_k & \mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_z, \hat{\psi}^2] & \boldsymbol{0}_{k\times k} & \mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{w}}_z] \end{bmatrix}, \end{aligned} \tag{2.16}$$

where $\boldsymbol{0}_k$ is a $k \times 1$ vector of zeros, $\boldsymbol{0}_{k\times k}$ a $k \times k$ matrix of zeros, and

$$\mathbb{V}\mathrm{ar}[\hat{\sigma}_g^2] = \frac{2(T-N)\sigma_g^4}{T^2}, \tag{2.17}$$

$$\mathbb{V}\mathrm{ar}[\hat{\mu}_g] = \frac{(T\psi^2 + T - 2)\sigma_g^2}{T(T-N-1)} \text{ if } T > N+1, \tag{2.18}$$

$$\mathbb{V}\mathrm{ar}[\hat{\psi}^2] = \frac{2T^2\psi^4 + 2(T-2)(2T\psi^2 + N - 1)}{(T-N-1)^2(T-N-3)} \text{ if } T > N+3, \tag{2.19}$$

$$\mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{w}}_g] = \frac{\sigma_g^2}{T-N-1} \boldsymbol{A} \text{ if } T > N+1, \tag{2.20}$$

$$\begin{aligned} \mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{w}}_z] &= \frac{T^2(T-N+1)}{(T-N)(T-N-1)^2(T-N-3)} \boldsymbol{L}\boldsymbol{w}_z \boldsymbol{w}_z^{\mathrm{T}} \boldsymbol{L}^{\mathrm{T}} \\ &\quad + \frac{T(T\psi^2 + T - 2)}{(T-N)(T-N-1)(T-N-3)} \boldsymbol{A} \text{ if } T > N+3, \end{aligned} \tag{2.21}$$

$$\mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_g, \hat{\mu}_g] = \frac{\sigma_g^2}{T-N-1}\boldsymbol{L}\boldsymbol{w}_z \text{ if } T > N+1, \tag{2.22}$$

$$\mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{w}}_z, \hat{\psi}^2] = \frac{2T(T\psi^2+T-2)}{(T-N-1)^2(T-N-3)}\boldsymbol{L}\boldsymbol{w}_z \text{ if } T > N+3. \tag{2.23}$$

The mean and covariance matrix of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2)$ are available in [6]. The mean and covariance matrix of $\boldsymbol{L}\hat{\boldsymbol{w}}_g$ and $\boldsymbol{L}\hat{\boldsymbol{w}}_z$ are available in [1] and [7]. However, the covariance matrix between $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2)$ and $(\boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ is new.

If, as in [2], we are interested in $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ instead of $\boldsymbol{L}\hat{\boldsymbol{w}}_z$, we can also use Theorem 2.2 to derive its mean and its covariance with $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{\eta}})$. These moments are more involved, and thus we present them with proofs in the next theorem.

Theorem 2.3. *The mean of $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ and its covariance with $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{\eta}})$ are*

$$\mathbb{E}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}\right] = \frac{T}{N-1}\,{}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)\boldsymbol{L}\boldsymbol{w}_z \text{ if } N > 2, \tag{2.24}$$

$$\mathbb{C}\mathrm{ov}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}, \left(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{\eta}}\right)\right] = \left[\boldsymbol{0}_k, \boldsymbol{0}_k, \mathbb{C}\mathrm{ov}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}, \hat{\psi}^2\right], \boldsymbol{0}_{k\times k}, \mathbb{V}\mathrm{ar}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}\right]\right], \tag{2.25}$$

where

$$\mathbb{C}\mathrm{ov}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}, \hat{\psi}^2\right] = \frac{T}{T-N-1}\left[1 - \left(1 + \frac{T\psi^2}{N-1}\right){}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)\right]\boldsymbol{L}\boldsymbol{w}_z \quad \text{if } N > 2 \text{ and } T > N+1, \tag{2.26}$$

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}}\right] = \frac{T}{(T-N)(N-3)}&\left[{}_1F_1\left(1; \frac{N-1}{2}; -\frac{T\psi^2}{2}\right) + \frac{T-N-1}{N-1}\right. \\ &\left.{}_1F_1\left(2; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)\right]\boldsymbol{A} + \frac{T^2}{N^2-1}\left[\frac{T-N-1}{T-N}\,{}_1F_1\left(2; \frac{N+3}{2}; -\frac{T\psi^2}{2}\right)\right. \\ &\left. - \frac{N+1}{N-1}\,{}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)^2\right]\boldsymbol{L}\boldsymbol{w}_z\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} \text{ if } N > 3.\end{aligned} \tag{2.27}$$

3. Asymptotic distributions

Another reason why the stochastic representations in Theorems 2.1 and 2.2 are valuable is that they allow us to easily derive asymptotic distributions of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$, which is what we study in this section.

3.1. *Standard asymptotic regime*

In the next theorem, we consider the standard asymptotic regime in which N is fixed while $T \to \infty$.

Theorem 3.1. *Let N be fixed while $T \to \infty$. Then, the asymptotic joint distribu-*

tion of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ *is*

$$\sqrt{T}\begin{bmatrix} \hat{\sigma}_g^2 - \sigma_g^2 \\ \hat{\mu}_g - \mu_g \\ \hat{\psi}^2 - \psi^2 \\ \boldsymbol{L}\hat{\boldsymbol{w}}_g - \boldsymbol{L}\boldsymbol{w}_g \\ \boldsymbol{L}\hat{\boldsymbol{w}}_z - \boldsymbol{L}\boldsymbol{w}_z \end{bmatrix} \xrightarrow{d} \mathcal{N}\left(\boldsymbol{0}_{2k+3}, \boldsymbol{V}_0\right), \tag{3.1}$$

where

$$\boldsymbol{V}_0 = \begin{bmatrix} 2\sigma_g^4 & 0 & 0 & \boldsymbol{0}_k^{\mathrm{T}} & \boldsymbol{0}_k^{\mathrm{T}} \\ 0 & \sigma_g^2(1+\psi^2) & 0 & \sigma_g^2\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} & \boldsymbol{0}_k^{\mathrm{T}} \\ 0 & 0 & 2\psi^2(2+\psi^2) & \boldsymbol{0}_k^{\mathrm{T}} & 2(1+\psi^2)\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} \\ \boldsymbol{0}_k & \sigma_g^2\boldsymbol{L}\boldsymbol{w}_z & \boldsymbol{0}_k & \sigma_g^2\boldsymbol{A} & \boldsymbol{0}_{k\times k} \\ \boldsymbol{0}_k & \boldsymbol{0}_k & 2(1+\psi^2)\boldsymbol{L}\boldsymbol{w}_z & \boldsymbol{0}_{k\times k} & \boldsymbol{L}\boldsymbol{w}_z\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} + (1+\psi^2)\boldsymbol{A} \end{bmatrix}. \tag{3.2}$$

Theorem 3.1 shows that $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ are consistent estimators when N is fixed. However, in practice N is not negligible relative to T and the fixed N asymptotic distribution may be an inaccurate approximation of the finite-sample distribution.

3.2. *High-dimensional asymptotic regime*

We now turn to the high-dimensional (or Kolmogorov) asymptotic regime in which $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in (0,1)$. This regime is commonly used in recent finance literature to study the high-dimensional distributional properties of estimated portfolios and their inputs; see, e.g., [8], [9], [10], [11], and [12].

Theorem 3.2. *Let* $N \to \infty$, $T \to \infty$, *and* $N/T \to \rho \in (0,1)$ *while* $(\sigma_g^2, \mu_g, \psi^2, \boldsymbol{L}\boldsymbol{w}_g, \boldsymbol{L}\boldsymbol{w}_z, \boldsymbol{A})$ *are fixed.*[e] *Then, the joint asymptotic distribution of* $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ *is*

$$\sqrt{T}\begin{bmatrix} \hat{\sigma}_g^2 - (1-\rho)\sigma_g^2 \\ \hat{\mu}_g - \mu_g \\ \hat{\psi}^2 - \frac{\psi^2+\rho}{1-\rho} \\ \boldsymbol{L}\hat{\boldsymbol{w}}_g - \boldsymbol{L}\boldsymbol{w}_g \\ \boldsymbol{L}\hat{\boldsymbol{w}}_z - \frac{1}{1-\rho}\boldsymbol{L}\boldsymbol{w}_z \end{bmatrix} \xrightarrow{d} \mathcal{N}\left(\boldsymbol{0}_{2k+3}, \boldsymbol{V}_\rho\right), \tag{3.3}$$

[e]We can also assume these quantities are not fixed but are bounded as $N \to \infty$, and replace them by their asymptotic limit in (3.3)–(3.4). Assuming they are fixed is sensible when the objective of using the asymptotic distribution is to approximate the exact finite-sample distribution.

where

$$\boldsymbol{V}_\rho = \begin{bmatrix} 2(1-\rho)\sigma_g^4 & 0 & 0 & \mathbf{0}_k^{\mathrm{T}} & \mathbf{0}_k^{\mathrm{T}} \\ 0 & \frac{\sigma_g^2(1+\psi^2)}{1-\rho} & 0 & \frac{\sigma_g^2}{1-\rho}\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} & \mathbf{0}_k^{\mathrm{T}} \\ 0 & 0 & \frac{2(\psi^4+2\psi^2+\rho)}{(1-\rho)^3} & \mathbf{0}_k^{\mathrm{T}} & \frac{2(1+\psi^2)}{(1-\rho)^3}\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} \\ \mathbf{0}_k & \frac{\sigma_g^2}{1-\rho}\boldsymbol{L}\boldsymbol{w}_z & \mathbf{0}_k & \frac{\sigma_g^2}{1-\rho}\boldsymbol{A} & \mathbf{0}_{k\times k} \\ \mathbf{0}_k & \mathbf{0}_k & \frac{2(1+\psi^2)}{(1-\rho)^3}\boldsymbol{L}\boldsymbol{w}_z & \mathbf{0}_{k\times k} & \frac{\boldsymbol{L}\boldsymbol{w}_z\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}+(1+\psi^2)\boldsymbol{A}}{(1-\rho)^3} \end{bmatrix}. \tag{3.4}$$

Note that we recover the fixed N asymptotic distribution by letting $\rho \to 0$.

Theorem 3.2 shows that $(\hat{\sigma}_g^2, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ are asymptotically biased in the high-dimensional regime. Moreover, except for the asymptotic variance of $\hat{\sigma}_g^2$, all the entries of the asymptotic covariance matrix $\boldsymbol{V}_\rho$ in (3.4) increase relative to the fixed N asymptotic regime, $\boldsymbol{V}_0$ in (3.2).

4. Conclusion

We derive a simple stochastic representation for the joint distribution of sample estimates of three scalar parameters and two vectors of portfolio weights that characterize the minimum-variance frontier. This stochastic representation is useful, among others, to draw observations efficiently, understand which parameters determine the distribution, derive exact finite-sample moments in closed-form, and come up with the asymptotic joint distribution in the standard and the high-dimensional regime. In the process, we extend related results in [2] and [3].

The stochastic representation we derive is of wide interest to researchers in the portfolio choice literature because it is both simple and general. Indeed, as we show, it covers many different sample portfolio strategies such as the global minimum-variance portfolio, the tangency portfolio, the two-fund and three-fund rules in [4], and the fully invested two-fund rule in [1]. Moreover, one can use our results to study not only linear functions of these portfolios, such as the out-of-sample portfolio mean return, but also nonlinear functions like the out-of-sample portfolio return variance.

Acknowledgments

Nathan Lassance thanks the Fonds de la Recherche Scientifique (F.R.S.-FNRS) for financial support under grant number J.0115.22.

Appendix A. Proofs of results

A.1. *Proof of Theorem 2.1*

Let $\tilde{\boldsymbol{P}}$ be an $N \times N$ orthonormal matrix whose first column is

$$\boldsymbol{\nu} := \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{1}_N. \tag{A.1}$$

Let

$$\boldsymbol{z} := \sqrt{T}\tilde{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\sqrt{T}\tilde{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}, \boldsymbol{I}_N), \tag{A.2}$$

$$\boldsymbol{W} := T\tilde{\boldsymbol{P}}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\hat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P} \sim \mathcal{W}_N(T-1, \boldsymbol{I}_N), \tag{A.3}$$

and they are mutually independent. Using $\boldsymbol{W}$ and $\boldsymbol{z}$, we can write $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \hat{\boldsymbol{w}}_g, \hat{\boldsymbol{w}}_z)$ as

$$\hat{\sigma}_g^2 = \frac{\sigma_g^2}{T(\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1)}, \tag{A.4}$$

$$\hat{\mu}_g = \frac{\sigma_g}{\sqrt{T}}\frac{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{z}}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}, \tag{A.5}$$

$$\hat{\psi}^2 = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{z} - \frac{(\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{z})^2}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}, \tag{A.6}$$

$$\hat{\boldsymbol{w}}_g = \frac{\sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}, \tag{A.7}$$

$$\hat{\boldsymbol{w}}_z = \sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\boldsymbol{W}^{-\frac{1}{2}}\left(\boldsymbol{I}_N - \frac{\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-\frac{1}{2}}}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right)\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{z}, \tag{A.8}$$

where $\boldsymbol{e}_1 := [1,\ \boldsymbol{0}_{N-1}^{\mathrm{T}}]^{\mathrm{T}}$.

Partition $\boldsymbol{W}$ and $\boldsymbol{W}^{-1}$ as

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} \\ \boldsymbol{W}_{21} & \boldsymbol{W}_{22} \end{bmatrix}, \tag{A.9}$$

$$\boldsymbol{W}^{-1} = \begin{bmatrix} \boldsymbol{W}^{11} & \boldsymbol{W}^{12} \\ \boldsymbol{W}^{21} & \boldsymbol{W}^{22} \end{bmatrix}, \tag{A.10}$$

where $\boldsymbol{W}_{11}$ and $\boldsymbol{W}^{11}$ are the (1,1) elements of $\boldsymbol{W}$ and $\boldsymbol{W}^{-1}$, respectively. Using Theorem 3.2.10 of [13], we have

$$v_1 := \boldsymbol{W}_{11\cdot 2} = \boldsymbol{W}_{11} - \boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{21} \sim \chi^2_{T-N}, \tag{A.11}$$

$$\boldsymbol{x} := -\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{W}_{21} \sim \mathcal{N}(\boldsymbol{0}_{N-1}, \boldsymbol{I}_{N-1}), \tag{A.12}$$

$$\boldsymbol{W}_{22} \sim \mathcal{W}_{N-1}(T-1, \boldsymbol{I}_{N-1}), \tag{A.13}$$

and they are mutually independent and independent of $\boldsymbol{z}$. Using the formula for the inverse of a partitioned matrix, we obtain

$$\boldsymbol{W}^{11} = \boldsymbol{W}_{11\cdot 2}^{-1} = \frac{1}{v_1}, \tag{A.14}$$

$$\boldsymbol{W}^{21} = -\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{21}\boldsymbol{W}_{11\cdot 2}^{-1} = \frac{\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x}}{v_1}, \tag{A.15}$$

$$\boldsymbol{W}^{22} = \boldsymbol{W}_{22}^{-1} + \boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{21}\boldsymbol{W}_{11\cdot 2}^{-1}\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1} = \boldsymbol{W}_{22}^{-1} + \frac{\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}_{22}^{-\frac{1}{2}}}{v_1}. \tag{A.16}$$

Let $\boldsymbol{P}$ be the last $N-1$ columns of $\tilde{\boldsymbol{P}}$. We define

$$\boldsymbol{y} := \begin{bmatrix} z_2 \\ \vdots \\ z_N \end{bmatrix} \sim \mathcal{N}\left(\sqrt{T}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}, \boldsymbol{I}_{N-1}\right), \tag{A.17}$$

and $u := \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} \sim \chi^2_{N-1}(T\psi^2)$. Using $\boldsymbol{y}$, we can write $\boldsymbol{z}$ as

$$\boldsymbol{z} = \boldsymbol{e}_1 z_1 + \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}, \tag{A.18}$$

and thus $\boldsymbol{W}^{-1}\boldsymbol{z}$ can be expressed as

$$\boldsymbol{W}^{-1}z = \boldsymbol{W}^{-1}\boldsymbol{e}_1 z_1 + \boldsymbol{W}^{-1}\begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}^{11} \\ \boldsymbol{W}^{21} \end{bmatrix} z_1 + \begin{bmatrix} \boldsymbol{W}^{12}\boldsymbol{y} \\ \boldsymbol{W}^{22}\boldsymbol{y} \end{bmatrix}. \tag{A.19}$$

This allows us to write

$$\hat{\mu}_g = \frac{\sigma_g}{\sqrt{T}}\left(\frac{\boldsymbol{W}^{11}z_1 + \boldsymbol{W}^{12}\boldsymbol{y}}{\boldsymbol{W}^{11}}\right) = \frac{\sigma_g}{\sqrt{T}}\left(z_1 + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{y}\right), \tag{A.20}$$

$$\hat{\sigma}_g^2 = \frac{\sigma_g^2 v_1}{T}, \tag{A.21}$$

$$\begin{aligned}
\hat{\psi}^2 &= \left(z_1\boldsymbol{e}_1 + \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}\right)^{\mathrm{T}} \boldsymbol{W}^{-\frac{1}{2}}\left(\boldsymbol{I}_N - \frac{\boldsymbol{W}^{-\frac{1}{2}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-\frac{1}{2}}}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right)\boldsymbol{W}^{-\frac{1}{2}}\left(z_1\boldsymbol{e}_1 + \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}\right) \\
&= \begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix}^{\mathrm{T}}\left(\boldsymbol{W}^{-1} - \frac{\boldsymbol{W}^{-1}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}}{\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{W}^{-1}\boldsymbol{e}_1}\right)\begin{bmatrix} 0 \\ \boldsymbol{y} \end{bmatrix} \\
&= \boldsymbol{y}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\boldsymbol{y},
\end{aligned} \tag{A.22}$$

$$\hat{\boldsymbol{w}}_g = \frac{\sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\begin{bmatrix} \boldsymbol{W}^{11} \\ \boldsymbol{W}^{21} \end{bmatrix}}{\boldsymbol{W}^{11}} = \sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\begin{bmatrix} 1 \\ \boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x} \end{bmatrix} = \boldsymbol{w}_g + \sigma_g\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x}, \tag{A.23}$$

$$\begin{aligned}
\hat{\boldsymbol{w}}_z &= \sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\left(\begin{bmatrix} \boldsymbol{W}^{11} \\ \boldsymbol{W}^{21} \end{bmatrix} z_1 + \begin{bmatrix} \boldsymbol{W}^{12}\boldsymbol{y} \\ \boldsymbol{W}^{22}\boldsymbol{y} \end{bmatrix} - \begin{bmatrix} \boldsymbol{W}^{11} \\ \boldsymbol{W}^{21} \end{bmatrix}\left(z_1 + \frac{\boldsymbol{W}^{12}}{\boldsymbol{W}^{11}}\boldsymbol{y}\right)\right) \\
&= \sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\begin{bmatrix} 0 \\ \boldsymbol{W}^{22}\boldsymbol{y} - \frac{\boldsymbol{W}^{21}\boldsymbol{W}^{12}\boldsymbol{y}}{\boldsymbol{W}^{11}} \end{bmatrix} \\
&= \sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\tilde{\boldsymbol{P}}\begin{bmatrix} 0 \\ \boldsymbol{W}_{22}^{-1}\boldsymbol{y} \end{bmatrix} \\
&= \sqrt{T}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{W}_{22}^{-1}\boldsymbol{y}.
\end{aligned} \tag{A.24}$$

Let $\tilde{\boldsymbol{R}} := [\tilde{\boldsymbol{z}},\ \boldsymbol{R}]$ be an $(N-1)\times(N-1)$ orthonormal matrix, where

$$\tilde{\boldsymbol{z}} := \frac{\boldsymbol{y}}{\sqrt{u}}, \tag{A.25}$$

and $\boldsymbol{R}$ is a null matrix when $N=2$. We define

$$\boldsymbol{H} := (\tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{R}})^{-1} = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix} \sim \mathcal{W}_{N-1}(T-1, \boldsymbol{I}_{N-1}), \tag{A.26}$$

where $\boldsymbol{H}_{1,1}$ is the (1,1) element of $\boldsymbol{H}$. Using Theorem 3.2.10 of [13], we have

$$v_2 := \boldsymbol{H}_{11\cdot 2} = \boldsymbol{H}_{11} - \boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21} \sim \chi^2_{T-N+1}, \tag{A.27}$$

$$\boldsymbol{x}_2 := -\boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{H}_{21} \sim \mathcal{N}(\boldsymbol{0}_{N-2}, \boldsymbol{I}_{N-2}), \tag{A.28}$$

$$\boldsymbol{H}_{22} \sim \mathcal{W}_{N-2}(T-1, \boldsymbol{I}_{N-2}), \tag{A.29}$$

and they are mutually independent and independent of v_1 and $\boldsymbol{x}$.

Using the formula for the inverse of a partitioned matrix again, we can show that

$$\tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{R}} = \boldsymbol{H}^{-1} = \begin{bmatrix} \boldsymbol{H}_{11\cdot 2}^{-1} & -\boldsymbol{H}_{11\cdot 2}^{-1}\boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1} \\ -\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{H}_{11\cdot 2}^{-1} & \boldsymbol{H}_{22}^{-1} + \boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{H}_{11\cdot 2}^{-1}\boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1} \end{bmatrix}. \tag{A.30}$$

This allows us to write

$$\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{R}}\tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{R}}\begin{bmatrix} \boldsymbol{H}_{11\cdot 2}^{-1} \\ -\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{H}_{11\cdot 2}^{-1} \end{bmatrix} = \frac{\tilde{\boldsymbol{R}}}{v_2}\begin{bmatrix} 1 \\ \boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{x}_2 \end{bmatrix} = \frac{\tilde{\boldsymbol{z}} + \boldsymbol{R}\boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{x}_2}{v_2}. \tag{A.31}$$

Let

$$\boldsymbol{q} := \tilde{\boldsymbol{R}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x}. \tag{A.32}$$

Conditional on $\boldsymbol{H}$, we have

$$\boldsymbol{q} \sim \mathcal{N}(\boldsymbol{0}_{N-1}, \boldsymbol{H}^{-1}). \tag{A.33}$$

The first element of q is

$$q_1 = \tilde{\boldsymbol{z}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-\frac{1}{2}}\boldsymbol{x} \sim \mathcal{N}(0, \tilde{\boldsymbol{z}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{z}}). \tag{A.34}$$

Letting

$$a := (\tilde{\boldsymbol{z}}^{\mathrm{T}}\boldsymbol{W}_{22}^{-1}\tilde{\boldsymbol{z}})^{-\frac{1}{2}}q_1 = \sqrt{v_2}q_1 \sim \mathcal{N}(0,1), \tag{A.35}$$

we can write $q_1 = a/\sqrt{v_2}$. Therefore, it remains to obtain a stochastic representation of $\tilde{\boldsymbol{q}} := [q_2, \ldots, q_{N-1}]^{\mathrm{T}}$.

Conditional on q_1, we have

$$\tilde{\boldsymbol{q}}|q_1 \sim \mathcal{N}\left(-\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}q_1, \boldsymbol{H}_{22}^{-1}\right), \tag{A.36}$$

and we can write

$$\tilde{\boldsymbol{q}} = -\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}q_1 + \boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{x}_1 = \frac{a}{\sqrt{v_2}}\boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{x}_2 + \boldsymbol{H}_{22}^{-\frac{1}{2}}\boldsymbol{x}_1, \tag{A.37}$$

where $\boldsymbol{x}_1 \sim \mathcal{N}(\boldsymbol{0}_{N-2}, \boldsymbol{I}_{N-2})$, and it is independent of $\boldsymbol{x}_2$ and $\boldsymbol{H}_{22}$. In addition, using Theorem 3.1 and Corollary 3.1 from [14], we can write

$$\boldsymbol{H}_{22}^{-\frac{1}{2}}[\boldsymbol{x}_1,\ \boldsymbol{x}_2] \stackrel{d}{=} [\tilde{\boldsymbol{x}}_1,\ \tilde{\boldsymbol{x}}_2]\boldsymbol{C}^{-1}, \tag{A.38}$$

where $\tilde{\boldsymbol{x}}_1 \sim \mathcal{N}(\boldsymbol{0}_{N-2}, \boldsymbol{I}_{N-2})$, $\tilde{\boldsymbol{x}}_2 \sim \mathcal{N}(\boldsymbol{0}_{N-2}, \boldsymbol{I}_{N-2})$, and $\boldsymbol{C}$ is a lower triangular matrix such that $\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}} \sim \mathcal{W}_2(T-N+3, \boldsymbol{I}_2)$, and $(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \boldsymbol{C})$ are mutually independent. Using the Bartlett decomposition (see [13], p.99), we can write

$$\boldsymbol{C}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{w_2}} & 0 \\ \frac{b}{\sqrt{w_1 w_2}} & \frac{1}{\sqrt{w_1}} \end{bmatrix}, \tag{A.39}$$

where $b \sim \mathcal{N}(0,1)$, $w_1 \sim \chi^2_{T-N+2}$, $w_2 \sim \chi^2_{T-N+3}$, and they are mutually independent. This allows us to write

$$\boldsymbol{H}_{22}^{-\frac{1}{2}} \boldsymbol{x}_1 \stackrel{d}{=} \frac{\tilde{\boldsymbol{x}}_1}{\sqrt{w_2}} + \frac{b\tilde{\boldsymbol{x}}_2}{\sqrt{w_1 w_2}}, \tag{A.40}$$

$$\boldsymbol{H}_{22}^{-\frac{1}{2}} \boldsymbol{x}_2 \stackrel{d}{=} \frac{\tilde{\boldsymbol{x}}_2}{\sqrt{w_1}}. \tag{A.41}$$

With all these random variables, we can write

$$\hat{\sigma}_g^2 = \frac{\sigma_g^2 v_1}{T}, \tag{A.42}$$

$$\hat{\mu}_g = \frac{\sigma_g}{\sqrt{T}} \left(z_1 + a\frac{\sqrt{u}}{\sqrt{v_2}} \right), \tag{A.43}$$

$$\hat{\psi}^2 = u\tilde{\boldsymbol{z}}^{\mathrm{T}} \boldsymbol{W}_{22}^{-1} \tilde{\boldsymbol{z}} = \frac{u}{v_2}, \tag{A.44}$$

$$\begin{aligned} \hat{\boldsymbol{w}}_g &= \boldsymbol{w}_g + \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\tilde{\boldsymbol{R}}\boldsymbol{q} \\ &= \boldsymbol{w}_g + \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}(\tilde{\boldsymbol{z}} q_1 + \boldsymbol{R}\tilde{\boldsymbol{q}}) \\ &= \boldsymbol{w}_g + \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left[\frac{a\boldsymbol{y}}{\sqrt{v_2 u}} + \boldsymbol{R}\left(\frac{a}{\sqrt{v_2}} \boldsymbol{H}_{22}^{-\frac{1}{2}} \boldsymbol{x}_2 + \boldsymbol{H}_{22}^{-\frac{1}{2}} \boldsymbol{x}_1 \right) \right] \\ &\stackrel{d}{=} \boldsymbol{w}_g + \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left[\frac{a\boldsymbol{y}}{\sqrt{v_2 u}} + \boldsymbol{R}\left(\frac{a\tilde{\boldsymbol{x}}_2}{\sqrt{v_2 w_1}} + \frac{\tilde{\boldsymbol{x}}_1}{\sqrt{w_2}} + \frac{b\tilde{\boldsymbol{x}}_2}{\sqrt{w_1 w_2}} \right) \right], \end{aligned} \tag{A.45}$$

$$\begin{aligned} \hat{\boldsymbol{w}}_z &= \sqrt{Tu} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \boldsymbol{W}_{22}^{-1} \tilde{\boldsymbol{z}} \\ &= \frac{\sqrt{Tu}}{v_2} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left(\tilde{\boldsymbol{z}} + \boldsymbol{R}\boldsymbol{H}_{22}^{-\frac{1}{2}} \boldsymbol{x}_2 \right) \\ &\stackrel{d}{=} \frac{\sqrt{T}}{v_2} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left(\boldsymbol{y} + \sqrt{u} \frac{\boldsymbol{R}\tilde{\boldsymbol{x}}_2}{\sqrt{w_1}} \right). \end{aligned} \tag{A.46}$$

Letting

$$\boldsymbol{t}_1 := \frac{\tilde{\boldsymbol{x}}_2}{\sqrt{w_1}} \sim \frac{t_{T-N+2}(\boldsymbol{I}_{N-2})}{\sqrt{T-N+2}}, \tag{A.47}$$

$$\boldsymbol{t}_2 := \frac{1}{\sqrt{w_2}} (\boldsymbol{I}_{N-2} + \boldsymbol{t}_1 \boldsymbol{t}_1^{\mathrm{T}})^{-\frac{1}{2}} (\boldsymbol{t}_1 b + \tilde{\boldsymbol{x}}_1) \sim \frac{t_{T-N+3}(\boldsymbol{I}_{N-2})}{\sqrt{T-N+3}}, \tag{A.48}$$

which are mutually independent, we can write

$$\hat{\boldsymbol{w}}_g \overset{d}{=} \boldsymbol{w}_g + \sigma_g \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left[\frac{a}{\sqrt{v_2 u}} \boldsymbol{y} + \boldsymbol{R} \left(\frac{a}{\sqrt{v_2}} \boldsymbol{t}_1 + (\boldsymbol{I}_{N-2} + \boldsymbol{t}_1 \boldsymbol{t}_1^{\mathrm{T}})^{\frac{1}{2}} \boldsymbol{t}_2 \right) \right], \tag{A.49}$$

$$\hat{\boldsymbol{w}}_z \overset{d}{=} \frac{\sqrt{T}}{v_2} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left(\boldsymbol{y} + \sqrt{u} \boldsymbol{R} \boldsymbol{t}_1 \right). \tag{A.50}$$

This completes the proof.

A.2. *Proof of Theorem 2.2*

From (A.45) and (A.46), we obtain

$$\boldsymbol{L}\hat{\boldsymbol{w}}_g = \boldsymbol{L}\boldsymbol{w}_g + \sigma_g \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left[\frac{a\boldsymbol{y}}{\sqrt{v_2 u}} + \boldsymbol{R} \left(\frac{a\tilde{\boldsymbol{x}}_2}{\sqrt{v_2 w_1}} + \frac{\tilde{\boldsymbol{x}}_1}{\sqrt{w_2}} + \frac{b\tilde{\boldsymbol{x}}_2}{\sqrt{w_1 w_2}} \right) \right], \tag{A.51}$$

$$\boldsymbol{L}\hat{\boldsymbol{w}}_z = \frac{\sqrt{T}}{v_2} \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P} \left(\boldsymbol{y} + \sqrt{u} \frac{\boldsymbol{R}\tilde{\boldsymbol{x}}_2}{\sqrt{w_1}} \right). \tag{A.52}$$

Let

$$\check{\boldsymbol{y}} := \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{y} \sim \mathcal{N}(\sqrt{T}\boldsymbol{L}\boldsymbol{w}_z, \boldsymbol{A}), \tag{A.53}$$

$$\check{\boldsymbol{x}}_1 := \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{R}\tilde{\boldsymbol{x}}_1 \sim \mathcal{N}(\boldsymbol{0}_k, \tilde{\boldsymbol{B}}), \tag{A.54}$$

$$\check{\boldsymbol{x}}_2 := \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{R}\tilde{\boldsymbol{x}}_2 \sim \mathcal{N}(\boldsymbol{0}_k, \tilde{\boldsymbol{B}}), \tag{A.55}$$

where

$$\boldsymbol{A} := \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{P}^{\mathrm{T}} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{L}^{\mathrm{T}} = \boldsymbol{L}\boldsymbol{Q}\boldsymbol{L}^{\mathrm{T}}, \tag{A.56}$$

$$\tilde{\boldsymbol{B}} := \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{L}^{\mathrm{T}} = \boldsymbol{A} - \frac{\boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{P}\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}} \boldsymbol{P}^{\mathrm{T}} \boldsymbol{\Sigma}^{-\frac{1}{2}} \boldsymbol{L}^{\mathrm{T}}}{u} = \boldsymbol{A} - \frac{\check{\boldsymbol{y}}\check{\boldsymbol{y}}^{\mathrm{T}}}{u}, \tag{A.57}$$

where $\boldsymbol{Q}$ is defined in (1.3). Depending on the matrix $\boldsymbol{L}$, the rank of $\boldsymbol{A}$ can be less than k, in which case $\boldsymbol{A}$ is not invertible. Suppose the rank of $\boldsymbol{A}$ is m, with $m \leq \min(k, N-1)$. We perform an eigen-decomposition of $\boldsymbol{A}$ as $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{D}$ is a diagonal matrix of the m nonzero eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{V}$ is an $k \times m$ matrix of the corresponding eigenvectors. Let

$$\tilde{\boldsymbol{y}} := \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{V}^{\mathrm{T}} \check{\boldsymbol{y}} \sim \mathcal{N}(\sqrt{T}\boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{V}^{\mathrm{T}} \boldsymbol{L}\boldsymbol{w}_z, \boldsymbol{I}_m). \tag{A.58}$$

Defining $u_0 := u - \tilde{\boldsymbol{y}}^{\mathrm{T}} \tilde{\boldsymbol{y}}$, we have

$$u_0 \sim \chi^2_{N-m-1}(T\psi^2 - T\boldsymbol{w}_z^{\mathrm{T}} \boldsymbol{L}^{\mathrm{T}} \boldsymbol{V}\boldsymbol{D}^{-1} \boldsymbol{V}^{\mathrm{T}} \boldsymbol{L}\boldsymbol{w}_z), \tag{A.59}$$

and it is independent of $\tilde{\boldsymbol{y}}$. When $m = N-1$, we set $u_0 = 0$. Note that we can set $\check{\boldsymbol{y}} = \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \tilde{\boldsymbol{y}}$ because

$$\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \tilde{\boldsymbol{y}} \sim \mathcal{N}(\sqrt{T}\boldsymbol{L}\boldsymbol{w}_z, \boldsymbol{A}). \tag{A.60}$$

The mean of $\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \tilde{\boldsymbol{y}}$ is obtained by using

$$\mathbb{E}[\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}} \tilde{\boldsymbol{y}}] = \sqrt{T}\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} \boldsymbol{L}\boldsymbol{w}_z = \sqrt{T}(\boldsymbol{I}_k - \tilde{\boldsymbol{V}}_0 \tilde{\boldsymbol{V}}_0^{\mathrm{T}})\boldsymbol{L}\boldsymbol{w}_z = \sqrt{T}\boldsymbol{L}\boldsymbol{w}_z, \tag{A.61}$$

where $\tilde{\boldsymbol{V}}_0$ is the $k \times (k-m)$ matrix of the eigenvectors associated with the zero eigenvalues of $\boldsymbol{A}$. The last equality holds because

$$\boldsymbol{L}\boldsymbol{w}_z = \boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu} \tag{A.62}$$

is a linear combination of the columns of $\boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}$, so it is in the span of $\boldsymbol{A}$ and we have $\tilde{\boldsymbol{V}}_0^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z = \boldsymbol{0}_{k-m}$.

Define

$$\boldsymbol{B} := \boldsymbol{I}_m - \frac{1-\sqrt{\frac{u_0}{u}}}{\tilde{\boldsymbol{y}}^{\mathrm{T}}\tilde{\boldsymbol{y}}}\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{T}}, \tag{A.63}$$

and it can be shown that

$$\begin{aligned}
\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}} &= \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\left(\boldsymbol{I}_m - \frac{1}{u}\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{T}}\right)\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}} \\
&= \boldsymbol{A} - \frac{\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{T}}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}}}{u} \\
&= \boldsymbol{A} - \frac{\check{\boldsymbol{y}}\check{\boldsymbol{y}}^{\mathrm{T}}}{u} = \tilde{\boldsymbol{B}}.
\end{aligned} \tag{A.64}$$

We can write $\check{\boldsymbol{x}}_1 = \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{B}\boldsymbol{q}_1$ and $\check{\boldsymbol{x}}_2 = \boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{B}\boldsymbol{q}_2$ with

$$\boldsymbol{q}_1 \sim \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m), \tag{A.65}$$

$$\boldsymbol{q}_2 \sim \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m), \tag{A.66}$$

and $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are mutually independent.

With these random variables, we can now write

$$\boldsymbol{L}\hat{\boldsymbol{w}}_g \stackrel{d}{=} \boldsymbol{L}\boldsymbol{w}_g + \sigma_g\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\left[\frac{a}{\sqrt{v_2 u}}\tilde{\boldsymbol{y}} + \boldsymbol{B}\left(\frac{\boldsymbol{q}_1}{\sqrt{w_2}} + \left(\frac{a}{\sqrt{v_2}} + \frac{b}{\sqrt{w_2}}\right)\frac{\boldsymbol{q}_2}{\sqrt{w_1}}\right)\right], \tag{A.67}$$

$$\boldsymbol{L}\hat{\boldsymbol{w}}_z \stackrel{d}{=} \frac{\sqrt{T}}{v_2}\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\left(\tilde{\boldsymbol{y}} + \sqrt{u}\frac{\boldsymbol{B}\boldsymbol{q}_2}{\sqrt{w_1}}\right). \tag{A.68}$$

This stochastic representation requires $3m+8$ random variables, and we can reduce this number to $3m+5$ by using

$$\boldsymbol{L}\hat{\boldsymbol{w}}_g \stackrel{d}{=} \boldsymbol{L}\boldsymbol{w}_g + \sigma_g\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\left[\frac{a}{\sqrt{v_2 u}}\tilde{\boldsymbol{y}} + \boldsymbol{B}\left(\frac{a}{\sqrt{v_2}}\tilde{\boldsymbol{t}}_1 + (\boldsymbol{I}_m + \tilde{\boldsymbol{t}}_1\tilde{\boldsymbol{t}}_1^{\mathrm{T}})^{\frac{1}{2}}\tilde{\boldsymbol{t}}_2\right)\right], \tag{A.69}$$

$$\boldsymbol{L}\hat{\boldsymbol{w}}_z \stackrel{d}{=} \frac{\sqrt{T}}{v_2}\boldsymbol{V}\boldsymbol{D}^{\frac{1}{2}}\left(\tilde{\boldsymbol{y}} + \sqrt{u}\boldsymbol{B}\tilde{\boldsymbol{t}}_1\right), \tag{A.70}$$

where

$$\tilde{\boldsymbol{t}}_1 := \frac{\boldsymbol{q}_2}{\sqrt{w_1}} \sim \frac{t_{T-N+2}(\boldsymbol{I}_m)}{\sqrt{T-N+2}}, \tag{A.71}$$

$$\tilde{\boldsymbol{t}}_2 := \frac{1}{\sqrt{w_2}}(\boldsymbol{I}_k + \tilde{\boldsymbol{t}}_1\tilde{\boldsymbol{t}}_1^{\mathrm{T}})^{-\frac{1}{2}}(\boldsymbol{q}_1 + b\tilde{\boldsymbol{t}}_1) \sim \frac{t_{T-N+3}(\boldsymbol{I}_m)}{\sqrt{T-N+3}}. \tag{A.72}$$

This completes the proof.

A.3. *Proof of Theorem 2.3*

Using Theorem 2.1, the mean of $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ is

$$\begin{aligned}\mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}] &= \sqrt{T}\boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\mathbb{E}\left[\frac{\boldsymbol{y}}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}}\right] \\ &= \sqrt{T}\boldsymbol{L}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\left[\frac{{}_1F_1\left(1;\frac{N+1}{2};-\frac{T\psi^2}{2}\right)}{N-1}\sqrt{T}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}\right] \\ &= \frac{T}{N-1}{}_1F_1\left(1;\frac{N+1}{2};-\frac{T\psi^2}{2}\right)\boldsymbol{L}\boldsymbol{w}_z, \end{aligned} \tag{A.73}$$

provided $N > 2$, where the second equality follows from Equation (EC.53) in [6].

The proof of $\mathbb{C}\mathrm{ov}\left[\boldsymbol{L}\hat{\boldsymbol{\eta}},\left(\hat{\sigma}_g^2,\hat{\mu}_g,\boldsymbol{L}\hat{\boldsymbol{w}}_g\right)\right] = (\boldsymbol{0}_k,\boldsymbol{0}_k,\boldsymbol{0}_{k\times k})$ is straightforward. Turning to the covariance between $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ and $\hat{\psi}^2$, we have

$$\mathbb{C}\mathrm{ov}[\boldsymbol{L}\hat{\boldsymbol{\eta}},\hat{\psi}^2] = \mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{w}}_z] - \mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}]\mathbb{E}[\hat{\psi}^2], \tag{A.74}$$

where $\mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{w}}_z]$, $\mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}]$, and $\mathbb{E}[\hat{\psi}^2]$ are available in (2.15) and (2.24), which yields (2.26). Finally, the covariance matrix of $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ is

$$\mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{\eta}}] = \boldsymbol{L}\mathbb{E}[\hat{\boldsymbol{\eta}}\hat{\boldsymbol{\eta}}^{\mathrm{T}}]\boldsymbol{L}^{\mathrm{T}} - \mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}]\mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}]^{\mathrm{T}}, \tag{A.75}$$

where $\mathbb{E}[\boldsymbol{L}\hat{\boldsymbol{\eta}}]$ is given by (2.24) and

$$\mathbb{E}[\hat{\boldsymbol{\eta}}\hat{\boldsymbol{\eta}}^{\mathrm{T}}] = T\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\mathbb{E}\left[\frac{1}{u^2}(\boldsymbol{y}+\sqrt{u}\boldsymbol{R}\boldsymbol{t}_1)(\boldsymbol{y}+\sqrt{u}\boldsymbol{R}\boldsymbol{t}_1)^{\mathrm{T}}\right]\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}. \tag{A.76}$$

Since $\boldsymbol{t}_1$ has zero mean and is independent of $\boldsymbol{y}$ and $\boldsymbol{R}$, we have

$$\mathbb{E}[\hat{\boldsymbol{\eta}}\hat{\boldsymbol{\eta}}^{\mathrm{T}}] = T\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\mathbb{E}\left[\frac{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}}{u^2}+\frac{\boldsymbol{R}\boldsymbol{t}_1\boldsymbol{t}_1^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}}{u}\right]\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}. \tag{A.77}$$

The matrix $\boldsymbol{R}$ is such that $\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} = \boldsymbol{I}_{N-1} - \boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}/u$, and thus

$$\mathbb{E}\left[\frac{\boldsymbol{R}\boldsymbol{t}_1\boldsymbol{t}_1^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}}{u}\right] = \frac{1}{T-N}\mathbb{E}\left[\frac{\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}}}{u}\right] = \frac{1}{T-N}\left(\mathbb{E}\left[\frac{1}{u}\right]\boldsymbol{I}_{N-1} - \mathbb{E}\left[\frac{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}}{u^2}\right]\right). \tag{A.78}$$

This means that $\mathbb{E}[\hat{\boldsymbol{\eta}}\hat{\boldsymbol{\eta}}^{\mathrm{T}}]$ becomes

$$\mathbb{E}[\hat{\boldsymbol{\eta}}\hat{\boldsymbol{\eta}}^{\mathrm{T}}] = \frac{T}{T-N}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}\left(\mathbb{E}\left[\frac{1}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}}\right]\boldsymbol{I}_{N-1} + (T-N-1)\mathbb{E}\left[\frac{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}}{(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y})^2}\right]\right)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}. \tag{A.79}$$

Using Lemma 3 and Equation (EC.55) in [6], we can show that for $N > 3$,

$$\mathbb{E}\left[\frac{1}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}}\right] = \frac{{}_1F_1\left(1;\frac{N-1}{2};-\frac{T\psi^2}{2}\right)}{N-3}, \tag{A.80}$$

$$\begin{aligned}\mathbb{E}\left[\frac{\boldsymbol{y}\boldsymbol{y}^{\mathrm{T}}}{(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y})^2}\right] &= \frac{{}_1F_1\left(2;\frac{N+1}{2};-\frac{T\psi^2}{2}\right)}{(N-1)(N-3)}\boldsymbol{I}_{N-1} \\ &\quad + \frac{{}_1F_1\left(2;\frac{N+3}{2};-\frac{T\psi^2}{2}\right)}{(N+1)(N-1)}T\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{P}. \end{aligned} \tag{A.81}$$

Plugging (A.80) and (A.81) into (A.79), the covariance matrix of $\boldsymbol{L}\hat{\boldsymbol{\eta}}$ becomes

$$\begin{aligned}\mathbb{V}\mathrm{ar}[\boldsymbol{L}\hat{\boldsymbol{\eta}}] = & \frac{T}{(N-3)(T-N)} {}_1F_1\left(1; \frac{N-1}{2}; -\frac{T\psi^2}{2}\right)\boldsymbol{A} \\ & + \frac{T(T-N-1)}{(N-1)(N-3)(T-N)} {}_1F_1\left(2; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)\boldsymbol{A} \\ & + \frac{T^2(T-N-1)}{(N^2-1)(T-N)} {}_1F_1\left(2; \frac{N+3}{2}; -\frac{T\psi^2}{2}\right)\boldsymbol{L}\boldsymbol{w}_z\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}} \\ & - \frac{T^2}{(N-1)^2} {}_1F_1\left(1; \frac{N+1}{2}; -\frac{T\psi^2}{2}\right)^2 \boldsymbol{L}\boldsymbol{w}_z\boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}, \end{aligned} \tag{A.82}$$

which yields (2.27) after simplifications. This completes the proof.

A.4. *Proof of Theorem 3.1*

Using the exact finite-sample formula for the mean and covariance matrix of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ in (2.15)–(2.16), we can directly obtain the asymptotic mean and covariance matrix in (3.1)–(3.2) when N is fixed while $T \to \infty$.

Therefore, we are left with proving that the asymptotic distribution is normal, which we show using Theorem 2.2. Let $\hat{a} := a/\sqrt{T}$, $\hat{\boldsymbol{y}} := \tilde{\boldsymbol{y}}/\sqrt{T}$, $\hat{z} := z/\sqrt{T}$, $\hat{u}_0 := u_0/T$, $\hat{v}_1 := v_1/T$, and $\hat{v}_2 := v_2/T$. Then, when N is fixed while $T \to \infty$, we have

$$\sqrt{T}\hat{a} \sim \mathcal{N}(0,1), \tag{A.83}$$

$$\sqrt{T}(\hat{\boldsymbol{y}} - \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z) \sim \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m), \tag{A.84}$$

$$\sqrt{T}(\hat{z} - \mu_g/\sigma_g) \sim \mathcal{N}(0,1), \tag{A.85}$$

$$\sqrt{T}(\hat{u}_0 - (\psi^2 - \boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z)) \tag{A.86}$$

$$\xrightarrow{d} \mathcal{N}(0, 4(\psi^2 - \boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z)), \tag{A.87}$$

$$\sqrt{T}(\hat{v}_1 - 1) \xrightarrow{d} \mathcal{N}(0,2), \tag{A.88}$$

$$\sqrt{T}(\hat{v}_2 - 1) \xrightarrow{d} \mathcal{N}(0,2), \tag{A.89}$$

$$\sqrt{T}\tilde{\boldsymbol{t}}_1 \xrightarrow{d} \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m), \tag{A.90}$$

$$\sqrt{T}\tilde{\boldsymbol{t}}_2 \xrightarrow{d} \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m). \tag{A.91}$$

Since the distribution of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ can be written as a function of $(\hat{a}, \hat{\boldsymbol{y}}, \hat{z}, \hat{u}_0, \hat{v}_1, \hat{v}_2, \tilde{\boldsymbol{t}}_1, \tilde{\boldsymbol{t}}_2)$, it is asymptotically normal from the delta method. This completes the proof.

A.5. *Proof of Theorem 3.2*

Using the exact finite-sample formula for the mean and covariance matrix of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ in (2.15)–(2.16), we can directly obtain the asymptotic mean and covariance matrix in (3.3)–(3.4) when $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in (0,1)$.

Therefore, we are left with proving that the asymptotic distribution is normal, which we show using Theorem 2.2. Let $\hat{a} := a/\sqrt{T}$, $\hat{\boldsymbol{y}} := \tilde{\boldsymbol{y}}/\sqrt{T}$, $\hat{z} := z/\sqrt{T}$, $\hat{u}_0 := u_0/T$, $\hat{v}_1 := v_1/T$, and $\hat{v}_2 := v_2/T$. Then, when $N \to \infty$, $T \to \infty$, and $N/T \to \rho \in (0,1)$, we have

$$\sqrt{T}\hat{a} \sim \mathcal{N}(0,1), \tag{A.92}$$

$$\sqrt{T}(\hat{\boldsymbol{y}} - \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z) \sim \mathcal{N}(\boldsymbol{0}_m, \boldsymbol{I}_m), \tag{A.93}$$

$$\sqrt{T}(\hat{z} - \mu_g/\sigma_g) \sim \mathcal{N}(0,1), \tag{A.94}$$

$$\begin{aligned}&\sqrt{T}(\hat{u}_0 - (\rho + \psi^2 - \boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z))\\ &\qquad \xrightarrow{d} \mathcal{N}(0, 2\rho + 4(\psi^2 - \boldsymbol{w}_z^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{D}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{w}_z)),\end{aligned} \tag{A.95}$$

$$\sqrt{T}(\hat{v}_1 - (1-\rho)) \xrightarrow{d} \mathcal{N}(0, 2(1-\rho)), \tag{A.96}$$

$$\sqrt{T}(\hat{v}_2 - (1-\rho)) \xrightarrow{d} \mathcal{N}(0, 2(1-\rho)), \tag{A.97}$$

$$\sqrt{T}\tilde{\boldsymbol{t}}_1 \xrightarrow{d} \mathcal{N}\left(\boldsymbol{0}_m, \frac{1}{1-\rho}\boldsymbol{I}_m\right), \tag{A.98}$$

$$\sqrt{T}\tilde{\boldsymbol{t}}_2 \xrightarrow{d} \mathcal{N}\left(\boldsymbol{0}_m, \frac{1}{1-\rho}\boldsymbol{I}_m\right). \tag{A.99}$$

Since the distribution of $(\hat{\sigma}_g^2, \hat{\mu}_g, \hat{\psi}^2, \boldsymbol{L}\hat{\boldsymbol{w}}_g, \boldsymbol{L}\hat{\boldsymbol{w}}_z)$ can be written as a function of $(\hat{a}, \hat{\boldsymbol{y}}, \hat{z}, \hat{u}_0, \hat{v}_1, \hat{v}_2, \tilde{\boldsymbol{t}}_1, \tilde{\boldsymbol{t}}_2)$, it is asymptotically normal from the delta method. This completes the proof.

References

[1] R. Kan, X. Wang and G. Zhou, Optimal portfolio choice with estimation risk: No risk-free asset case, *Manage. Sci.* **68** (2022) 2047–2068.

[2] T. Bodnar, H. Dette, N. Parolya and E. Thorstén, Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions, *Random Matrices: Theory Appl.* **11** (2022).

[3] T. Bodnar, H. Dette, N. Parolya and E. Thorstén, Corrigendum to "Sampling distributions of optimal portfolio weights and characteristics in low and large dimensions", *arXiv* **1908.04243** (2023).

[4] R. Kan and G. Zhou, Optimal portfolio choice with parameter uncertainty, *J. Financial Quant. Anal.* **42** (2007) 621–656.

[5] N. Lassance, A. Martín-Utrera and M. Simaan, The risk of expected utility under parameter uncertainty, *Manage. Sci.*, forthcoming.

[6] R. Kan and D. Smith, The distribution of the sample minimum-variance frontier, *Manage. Sci.* **54** (2008) 1364–1380.

[7] Y. Okhrin and W. Schmid, Distributional properties of portfolio weights, *J. Econometrics* **134** (2006) 235–256.

[8] O. Ledoit and M. Wolf, Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks, *Rev. Financ. Stud.* **30** (2017) 4349–4388.

[9] T. Bodnar, N. Parolya and W. Schmid, Estimation of the global minimum variance portfolio in high dimensions, *Eur. J. Oper. Res.* **266** (2018) 371–390.

[10] M. Ao, L. Yingying and X. Zheng, Approaching mean-variance efficiency for large portfolios, *Rev. Financ. Stud.* **32** (2019) 2890–2919.

[11] T. Bodnar, Y. Okhrin and N. Parolya, Optimal shrinkage-based portfolio selection in high dimensions, *J. Bus. Econ. Stat.* **41** (2022) 140–156.

[12] R. Kan, X. Wang and X. Zheng, In-sample and out-of-sample Sharpe ratios of multi-factor asset pricing models, *SSRN working paper* (2022).

[13] R. Muirhead, *Aspects of Multivariate Statistical Theory* (John Wiley & Sons, New Jersey, 1982).

[14] J. Dickey, Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution, *Ann. Math. Stat.* **38** (1967) 511–518.