Sample Recycling via Likelihood Ratio for Nested Simulation

Kun Zhang*[1], Mingbin Feng†[2], Guangwu Liu‡[3], and Shiyu Wang§[3]

[1]Institute of Statistics and Big Data
Renmin University of China
Beijing, China

[2]Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, ON, Canada

[3]Department of Management Sciences
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong, China

June 1, 2021

Abstract

NEED TO CHANGE LATER. We introduce the stochastic mesh approach for portfolio risk measurement under the nested setting, where a two-level simulation proceeds. The outer-level simulation generates financial scenarios, whereas the inner-level estimates future portfolio values in each scenario and then compute portfolio losses. The stochastic mesh approach is well-known in pricing American option, and we use it to provide an unbiased and strong consistent estimator of portfolio loss. In this paper, we focus on three portfolio risk measures: squared tracking error, expected excess loss and the probability of a large loss, and the corresponding stochastic mesh estimators are constructed. On one hand, we provide central limit theorems related to the

*kunzhang@ruc.edu.cn
†ben.feng@uwaterloo.ca
‡msgw.liu@cityu.edu.hk
§shiyuwang7-c@my.cityu.edu.hk

three portfolio risk measures, and then construct asymptotically valid confidence intervals for the portfolio risk. On the other hand, the mean squared error (MSE) of the stochastic mesh estimators of portfolio risk is analyzed, and its convergence rate is of order Γ^{-1}, where Γ measures computational effort. Specifically, the method we introduce is feasible for portfolios consisting of path-dependent derivatives. Numerical study is consistent with the theory we present.

Key words: stochastic mesh; risk measurement; confidence intervals; nested simulation

1 Introduction

In this study we propose, analyze, and test an efficient simulation procedure for estimating

$$\rho = \mathbb{E}\left[g(\mathbb{E}\left[H|\boldsymbol{S}_\tau\right])\right] \tag{1.1}$$

where $\boldsymbol{S}_\tau$ is a random vector in $\mathbb{R}^d$ for some $d \geq 1$, H is a random variable in $\mathbb{R}$ whose source of randomness is characterized by $\boldsymbol{S}_\tau$, and $g(\cdot)$ is a given nonlinear function. Gets to the point in the first sentence. But a different set of notations will be used later.

Such *nested estimation* problem (Hong et al., 2017) is common and important in portfolio risk measurement. Estimating common risk measures such as exceedance probability, conditional value-at-risk (CVaR), or mean squared error (MSE) of a portfolio of complex derivatives can all be formulated as (1.1). Consider a portfolio of complex derivatives whose values are affected by different risk factors such as equity returns, interest rates, exchange rates, etc. Projection of the underlying risk factors and valuation of the derivatives both require Monte Carlo simulation, resulting in a *nested simulation* (Gordy and Juneja, 2010; Broadie et al., 2011) procedure, which is also known as two-level and stochastic-on-stochastic simulations. In the outer level, one simulates a number, say n_{out}, of scenarios of financial risk factors $\boldsymbol{S}_\tau$ at a future time τ, i.e., the *risk horizon*, when risk measurement is needed. In the inner level of each scenario $\boldsymbol{S}_\tau$, a number, say n_{in}, of sample paths of the underlying assets of the derivatives are simulated to estimate the portfolio loss, i.e., $\mathbb{E}[H|\boldsymbol{S}_\tau]$, by an unbiased sample average estimator. Depending on the complexity of the risk factor models, derivative payoffs, and the size of the portfolio, each inner simulation can be quite time-consuming. In a nested simulation procedure, a total of $\Gamma = n_{out} \times n_{in}$ inner-level simulations are required, which can be a prohibitively excessive computational burden. In theory, the risk estimator in a nested simulation procedure converges to the true risk measure as the numbers of outer and inner simulations grows. However, in practice, it may require unbearably large computations to achieve satisfactory accuracy.

Much research has been done to make nested simulation more practical. One such line of research focuses on intelligent ways to allocate a fixed simulation budget Γ so that the resulting risk estimator converges quickly. Lee (1998), Lee and Glynn (2003), and Gordy and Juneja (2010) analyze the nested simulation estimator and demonstrate that the optimal asymptotic mean squared error (MSE) of the standard nested risk estimator diminishes at rate $\Gamma^{-2/3}$ under some assumptions. Gordy and Juneja (2010) also shows that this optimal rate of convergence is achieved when $n_{out} = \mathcal{O}(\Gamma^{2/3})$ and $n_{in} = \mathcal{O}(\Gamma^{1/3})$. Along the same line of research, Broadie et al. (2011) proposes algorithms that sequentially allocate computational effort to inner-level simulations for estimating the probabilities of large portfolio losses. The MSE of the resulting risk estimator is shown to have a faster rate of convergence, i.e., $\Gamma^{-4/5+\varepsilon}$ for any $\varepsilon > 0$. In addition, Lan et al. (2010) use ranking-and-selection techniques to improve the efficiency of the inner estimation.Question: Does this work fit in this "budget allocation" category? In a nested simulation procedure, the conditional expectation for each scenario, i.e., $\mathbb{E}[H|\boldsymbol{S}_\tau]$, is estimated by the sample average of inner simulations of that scenario, without considering those of other scenarios. While the sample average is an unbiased estimator, such exclusivity is a wasteful use of simulation efforts.

Another relevant line of research aims to find alternatives of standard nested simulation, where the conditional expectation $\mathbb{E}[H|\boldsymbol{S}_\tau]$ can be estimated efficiently without inner simulations. For example, least-square Monte Carlo (LSMC) (Longstaff and Schwartz, 2001; Tsitsiklis and Van Roy, 2001) is a quintessential parametric approach for pricing American options, where conditional expectation $\mathbb{E}[H|\boldsymbol{S}_\tau]$ is approximated by a regression model. See also Carriere (1996) for a general discussions of nonparametric regression techniques in Monte Carlo simulation. Broadie et al. (2015) applies this LSMC approach in nested estimation of financial risk and show that the MSE of the resulting risk estimator achieve a fast convergence rate of Γ^{-1} to an asymptotic squared bias level instead of zero. The risk estimator in a LSMC procedure is biased in general, and the difficulty to select appropriate basis functions in the regression model hinders the use of LSMC in practical application. In a similar spirit, Liu and Staum (2010) considers a metamodeling approach that estimates $\mathbb{E}[H|\boldsymbol{S}_\tau]$ by a stochastic kriging modelcitation for stochastic kriging (**?**)watson1964kernel, which is also biased in general. Also, implementation of stochastic kriging is not trivial and may be prone to numerical instabilitycitation for numerical instability in kriging/stochastic kriging. Hong et al. (2017) proposes a kernel smoothing approach, which estimates $\mathbb{E}[H|\boldsymbol{S}_\tau]$ by the well-known Nadaraya-Watson kernel estimator (citation for this kernel estimator) (**??**). The MSE of the resulting risk estimator achieves a convergence rate of $\Gamma^{-\min\{1,4/(d+2)\}}$, where d is the problem dimension. Is the Nadaraya-Watson kernel estimator biased in general? (Asymptotically unbiased) These approaches first pool simulation efforts in a pilot experiment, either for calibrating a regression model/metamodel or for constructing a kernel estimator, then use the resulting model to approximate the conditional expectation $\mathbb{E}[H|\boldsymbol{S}_\tau]$. But these approaches produce biased estimators so their MSEs do not converge to zero in general. Also, model selections, model calibrations, and detailed implementations of these methods may be complicated and difficult.

In this article we study a novel approach, called the *green nested simulation (GNS)* procedure, that has the advantages of both nested simulation and its alternatives without suffering their difficulties. In essence, GNS pools and reuses the same set of simulations to estimate the conditional expectation $\mathbb{E}[H|\boldsymbol{S}_\tau]$ for different scenarios; the estimation is unbiased as the simulation outputs are weighted by appropriate likelihood ratios. GNS is inspired by the recently proposed green simulation and likelihood ratio metamodeling, which aims to improve simulation efficiency by recycling and reusing simulation outputs.citations for green simulations and likelihood ratio metamodeling GNS is also related to the stochastic mesh approach for pricing American options (Broadie et al., 2000; Broadie and Glasserman, 2004; Avramidis and Hyden, 1999; Avramidis and Matzinger, 2004). See add citation of QMC on stochastic mesh for Bermudan-type options. With appropriate likelihood ratio weights, stochastic mesh uses an unbiased estimator for the continuation value and subsequently produce an asymptotically unbiased estimator of American options that converges quickly. GNS and the stochastic mesh are similar mathematically in that both approaches use likelihood ratio weights to produce unbiased estimators for some conditional expectations. But the two approaches are developed to tackle different problems, serve different purposes, and are applied in different contexts. GNS aims to solve a nested estimation problem efficiently, while stochastic mesh aims to solve a dynamic programming problem. In addition, GNS considers a risk measurement problem at a fixed risk horizon, while stochastic mesh method was applied to American option pricing that has multiple potential exercise times. I feel this paragraph is rather weak. I am trying to differentiate GNS from stochastic mesh. The previous version says "GNS is an application of stochastic mesh to portfolio risk measure", which seems to be an understatement to me. Most importantly, our study includes further theoretical analysis than existing studies on stochastic mesh and nested simulation.

A main contribution of our study is the in-depth analysis of the theoretical properties of the proposed green simulation procedure. Under some assumptions, we show that:

1. The green simulation estimator for the conditional expectation is an unbiased and strongly consistent for a given scenario. Moreover, when the scenario is stochastic, we show that this estimator converges almost surely to the target loss random variable. This is an extension of the theoretical properties of importance sampling estimators.

2. The asymptotic bias, variance, and MSE of the resulting risk estimator all converge to zero at rate $\mathcal{O}(\Gamma^{-1})$. This convergence rate is faster than that of nested stimulation with optimal allocation and that of kernel approach. Also, unlike LSMC and other metamodeling-based approaches, the asymptotic bias of our GNS procedure vanishes. Most importantly, $\mathcal{O}(\Gamma^{-1})$ is the same fast convergence rate as a non-nested Monte Carlo simulation.

3. The risk estimator is asymptotically normally distributed with a classical convergence rate $\mathcal{O}(\Gamma^{-1/2})$, i.e., a central limit theorem (CLT) result is established. To the best of our knowledge, this is the first CLT result [with convergence rate $\mathcal{O}(\Gamma^{-1/2})$] for nested estimation problems.

4. The consistent estimator of the asymptotic variance in CLT is proposed and analyzed. Combining with the previous CLT results, valid confidence intervals are established. To the best of our knowledge, this is the first theoretically valid confidence interval for nested estimation problems.

Extensive numerical experiments are conducted and they provide supportive evidences to our theoretical analysis. Elaborate this paragraph when experiments are completed.

The rest of this chapter is organized as follows. The problem statement and general mathematical framework are given in Section 2. Sections **??** and **??** present the main asymptotic analyses: Section **??** analyzes the convergence of the green loss estimator to the conditional expectation random variable and Section **??** analyzes the asymptotic bias, variance, MSE, as well as the CLT and valid confidence interval of the portfolio risk estimator. Numerical experiments are summarized in Section **??**, followed by conclusions in Section **??**. Technical proofs and auxiliary discussions are provided in the appendices.

2 A Green Simulation Approach

2.1 Settings

Let S_t be a vector of risk factors, which may be the values of equities, bonds, interest rates, exchange rates, etc., at any time $t \geq 0$. Suppose these risk factors are modeled by a stochastic process that is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a natural filtration $\mathcal{F}_t$ governing its evolution. Consider a portfolio of financial instruments, which may include stocks, bonds, and derivatives whose values are affected by the risk factors. Let $t = 0$ be the current time where the initial risk factor values S_0 are known and $T > 0$ be the maximum maturity of all the instruments in the portfolio. We are interested in estimating the some risk measures of the portfolio loss at a fixed future time $\tau \in (0, T)$ called the *risk horizon*, due to the stochasticity of the risk factors. Specifically, let V_t be the portfolio value, i.e., the total values of the instruments in the portfolio, at any time $t \geq 0$. Under the risk-neural pricing framework (???REFERENCE NEEDED), V_t is the risk-neutrally discounted portfolio payoff that is measurable to $\mathcal{F}_t$. Then the portfolio loss at time t is $L_t = V_0 - V_t$, which is a random variable at time $t > 0$. We are interested in estimating some risk measures of this loss random variable.

For convenience, we denote the risk factors *up to* τ by $\boldsymbol{S}_\tau = \{S_t : t \in [0.\tau]\}$, which are referred to as the *front paths*. These are often the outer-level sample paths, or the outer scenarios, in a nested simulation. Similarly, risk factors *exceeding* τ are denoted by $\widetilde{\boldsymbol{S}}_{\tau^+} = \{S_t : t \in (\tau, T]\}$ and are referred to as the *back paths*. These are often the inner-level sample paths in a nested simulation. The different notations for front and back paths are chosen purposely: Firstly, in a risk measurement experiment they are often simulated under different measures, i.e., the front paths are simulated under the real-world measure which the back paths are simulated under the risk-neutral measure. Secondly, in parts of our analysis and numerical experiments the front and back paths may be simulated independently. Mathematically, the portfolio loss random variable is given by

$$L_\tau = L(\boldsymbol{S}_\tau) = V_0 - V_\tau = \mathbb{E}[H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})|\mathcal{F}_\tau] = \mathbb{E}[H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})|\boldsymbol{S}_\tau], \tag{2.1}$$

where $H(\cdot)$ is a real-valued function of the discounted portfolio loss given a sample path and the expectation is taken with respect to the conditional distribution of $\widetilde{\boldsymbol{S}}_{\tau^+}$ given $\boldsymbol{S}_\tau$. For a simulated sample path $(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})$, we call $H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})$ the *simulation output*. As alluded in the notations, the analyses in this article are applicable to non-Markovian risk factors and path-dependent portfolio payoffs. Then, for a given *risk function* $g : \mathbb{R} \mapsto \mathbb{R}$, the risk measure of interest is

$$\rho := \mathbb{E}[g(L(\boldsymbol{S}_\tau))]. \tag{2.2}$$

Some important risk measures in practice can be written as (2.2): If $g(\cdot)$ is a quadratic function, i.e., $g(x) = (x-u)^2$, then $\mathbb{E}[(L(\boldsymbol{S}_\tau)-u)^2]$ measures the squared tracking error of the portfolio loss relative to a target u. If $g(\cdot)$ is a hockey-stick function, i.e., $g(x) = (x-u)^+ = \max\{x-u, 0\}$, then $\mathbb{E}[(L(\boldsymbol{S}_\tau)-u)^+]$ measures the excess loss beyond a threshold u. If $g(\cdot)$ is an indicator function, i.e., $g(x) = \mathbb{1}\{x \geq u\}$, then $\mathbb{E}[\mathbb{1}\{L(\boldsymbol{S}_\tau) \geq u\}] = \mathbb{P}\{L(\boldsymbol{S}_\tau) \geq u\}$ measures the exceedance probability that the portfolio loss is larger than a threshold u. These three classes of risk functions also entail other important risk measures in enterprise risk management such as the mean squared error (MSE), the Conditional Value-at-Risk (CVaR)[1], and the Value-at-Risk (VaR). These three types of risk functions may be used to approximate more general risk functions, such as those with a finite number of non-differentiable or discontinuous points (see discussions in Hong et al., 2017). Without loss of generality, hereinafter we consider $u = 0$ and all analysis can be trivially extended to any threshold $u \in \mathbb{R}$ with a change of variable.

Nested simulation is often used to estimate (2.2) when analytical formula for the portfolio loss $L(\boldsymbol{s}_\tau)$ for given $\boldsymbol{s}_\tau$ is not available. Specifically, in a standard nested simulation procedure, one first simulate n independent front paths, $\boldsymbol{S}_\tau^{(1)}, \ldots, \boldsymbol{S}_\tau^{(n)}$, under the real-world measure. For each of the outer scenario $\boldsymbol{S}_\tau^{(i)}$, one further simulates m back paths, $\widetilde{\boldsymbol{S}}_{\tau^+}^{(i1)}, \ldots, \widetilde{\boldsymbol{S}}_{\tau^+}^{(im)}$ under the risk-neutral measure, or the *inner simulations*, to estimate $L(\boldsymbol{S}_\tau^{(i)})$ by $L_m^{NS}(\boldsymbol{S}_\tau^{(i)}) = \frac{1}{m}\sum_{j=1}^m H(\boldsymbol{S}_\tau^{(i)}, \widetilde{\boldsymbol{S}}_{\tau^+}^{(ij)})$. Then the empirical distribution of inner estimators $L_m^{NS}(\boldsymbol{S}_\tau^{(1)}), \ldots, L_m^{NS}(\boldsymbol{S}_\tau^{(n)})$ can be used to estimate the desired risk measure by, say $\rho_{mn}^{NS} = \frac{1}{n}\sum_{i=1}^n g(L_m^{NS}(\boldsymbol{S}_\tau^{(i)}))$. Nested simulation is a wasteful use of computations because each inner estimator $L_m^{NS}(\boldsymbol{S}_\tau^{(i)})$ only uses the m inner stimulation outputs associated with the ith front path, ignoring all other $(n-1)m$ outputs in other front paths. In the next section we propose an efficient green simulation procedure that makes use of all available simulation outputs for every inner estimator.

2.2 Green Nested Simulation (GNS) via Likelihood Ratios

Let $\widetilde{f}(\widetilde{\boldsymbol{s}}_{\tau^+})$ be a *sampling density* of the back paths and $f(\widetilde{\boldsymbol{s}}_{\tau^+}|\boldsymbol{s}_\tau)$ be the conditional density of back paths given a front path $\boldsymbol{s}_\tau$. We assume that the users can sample from $\widetilde{f}(\widetilde{\boldsymbol{s}}_{\tau^+})$ and can calculate

[1] Also known as the expected shortfall (ES) and conditional tail expectation (CTE).

values for both $\widetilde{f}(\widetilde{\boldsymbol{s}}_{\tau+})$ and $f(\widetilde{\boldsymbol{s}}_{\tau+}|\boldsymbol{s}_\tau)$. Then, similar to importance sampling the conditional loss (2.1) can be written as

$$L(\boldsymbol{s}_\tau) = \mathbb{E}\left[H(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{S}}_{\tau+})\frac{f(\widetilde{\boldsymbol{S}}_{\tau+}|\boldsymbol{s}_\tau)}{\widetilde{f}(\widetilde{\boldsymbol{S}}_{\tau+})}\right] =: \mathbb{E}\left[\widehat{H}(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{S}}_{\tau+})\right], \quad \widetilde{\boldsymbol{S}}_{\tau+} \sim \widetilde{f}, \tag{2.3}$$

where the sampling density satisfies that $H(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{s}}_{\tau+})f(\widetilde{\boldsymbol{s}}_{\tau+}|\boldsymbol{s}_\tau) = 0$ whenever $\widetilde{f}(\widetilde{\boldsymbol{s}}_{\tau+}) = 0$. The shorthand notation $\widehat{H}(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{s}}_{\tau+}) := H(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{s}}_{\tau+})\frac{f(\widetilde{\boldsymbol{s}}_{\tau+}|\boldsymbol{s}_\tau)}{\widetilde{f}(\widetilde{\boldsymbol{s}}_{\tau+})}$ is used to denote the likelihood-ratio-weighted simulation output. Despite mathematical similarities, our proposed method has a different objective than importance sampling. As discussed in Feng and Staum (2017), green simulation differs from importance sampling in use (2.3) to reuse simulation outputs while important sampling is often used for variance reduction.

[Examples of process with known density functions]

Inspired by (2.3), we propose the following green nested simulation (GNS) procedure:

1. Simulate m i.i.d. back paths, $\widetilde{\boldsymbol{S}}_{\tau+}^{(1)}, \ldots, \widetilde{\boldsymbol{S}}_{\tau+}^{(m)}$; these are the "inner simulations".

2. Simulate n i.i.d. front paths, $\boldsymbol{S}_\tau^{(1)}, \ldots, \boldsymbol{S}_\tau^{(n)}$; these are the "outer simulations". For each front path $\boldsymbol{S}_\tau^{(i)}$, $i = 1, \ldots, n$, estimate the portfolio loss $L(\boldsymbol{S}_\tau^{(i)})$ by

$$L_m(\boldsymbol{S}_\tau^{(i)}) = \frac{1}{m}\sum_{j=1}^m H(\boldsymbol{S}_\tau^{(i)}, \widetilde{\boldsymbol{S}}_{\tau+}^{(j)})\frac{f(\widetilde{\boldsymbol{S}}_{\tau+}^{(j)}|\boldsymbol{S}_\tau^{(i)})}{\widetilde{f}(\widetilde{\boldsymbol{S}}_{\tau+}^{(j)})} = \frac{1}{m}\sum_{j=1}^m \widehat{H}(\boldsymbol{S}_\tau^{(i)}, \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}). \tag{2.4}$$

3. Estimate the risk measure ρ in (1.1) by

$$\rho_{mn} = \frac{1}{n}\sum_{i=1}^n g(L_m(\boldsymbol{S}_\tau^{(i)})). \tag{2.5}$$

For notational convenience, where no confusion will arise we write simply L, L_m, and $\widehat{H}$ in places for $L(\boldsymbol{S}_\tau)$, $L_m(\boldsymbol{S}_\tau)$, and $\widehat{H}(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau+})$ respectively. For simulated sample paths we will use shorthand notations $L^{(i)}$, $L_m^{(i)}$, and $\widehat{H}^{(ij)}$ for $L(\boldsymbol{S}_\tau^{(i)})$, $L_m(\boldsymbol{S}_\tau^{(i)})$, and $\widehat{H}(\boldsymbol{S}_\tau^{(i)}, \widetilde{\boldsymbol{S}}_{\tau+}^{(j)})$, respectively. For example, we may write $\rho_{mn} = \frac{1}{n}\sum_{i=1}^n g(L_m^{(i)}) = \frac{1}{n}\sum_{i=1}^n g(\frac{1}{m}\sum_{j=1}^m \widehat{H}^{(ij)})$.

A graphical illustration of the GNS procedure is shown in Figure 1. The main ideas of the GNS procedure is that a back path that is not simulated from the conditional density $f(\cdot|\boldsymbol{s}_\tau)$ can also be used to estimate a conditional loss $L(\boldsymbol{s}_\tau)$. This is similar to the stochastic mesh approach for pricing American option. Interested readers are referred to Broadie et al. (2000), Section 8.5 of Glasserman (2013), and Liu and Hong (2009) for further details.

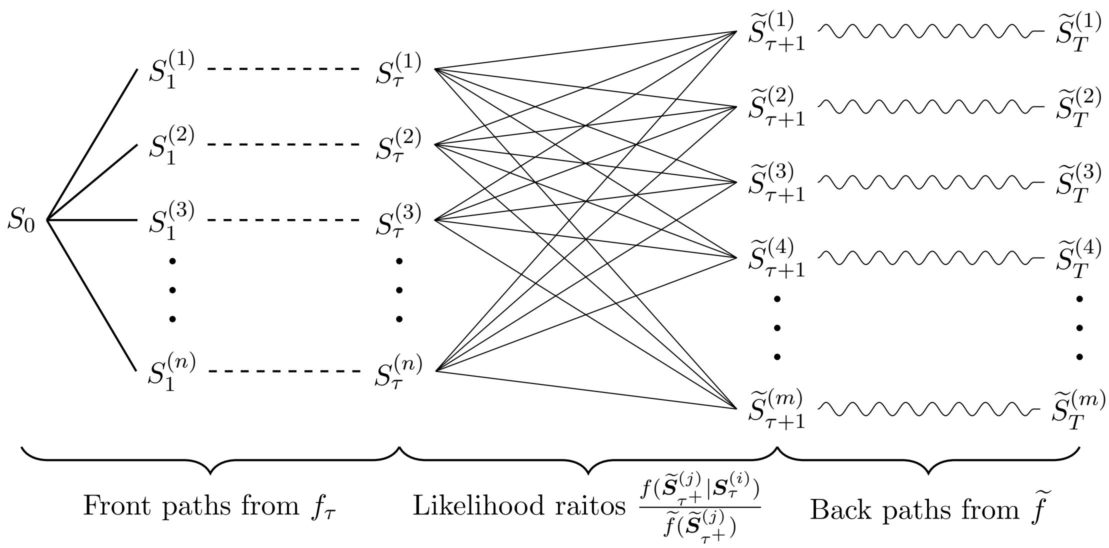

Figure 1: Illustration of reusing sample paths in green estimators L_m and ρ_{mn}.

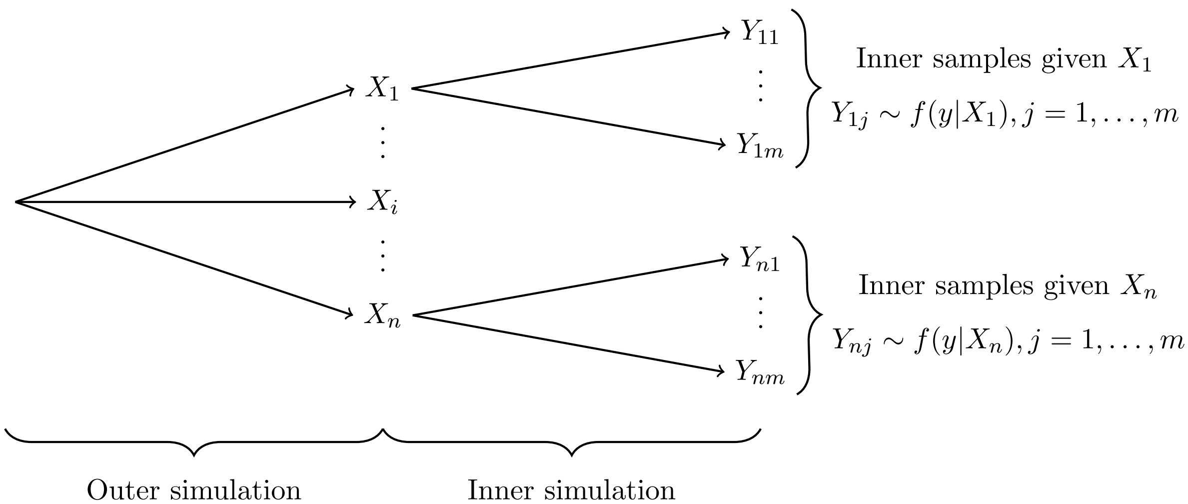

Figure 2: Illustration of reusing sample paths in green estimators L_m and ρ_{mn}.

The GNS procedure is flexible, as it allows the user to choose a sampling density $\widetilde{f}$ that may be different from any of the condition densities. The GNS procedure is efficient, as it recycles and reuses the same m back paths to estimate all n portfolio losses. In contrast, in a nested simulation procedure each portfolio loss $L(\boldsymbol{S}_\tau^{(i)})$ is estimated by its own set of inner sample paths. Also, with similar computations to a nested simulation procedure, the GNS procedure can afford $m = n_{in} \times n_{out}$

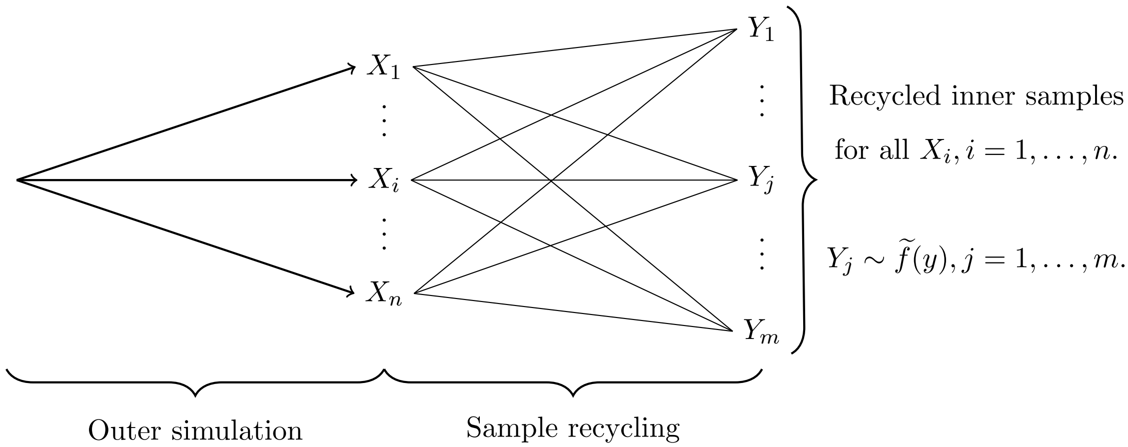

Figure 3: Illustration of reusing sample paths in green estimators L_m and ρ_{mn}.

back paths, so the latter expected to produce more accurate estimates of the portfolio losses and of the risk measure. Moreover, the likelihood ratios do not depend on the payoffs structure, i.e., the function $H(\cdot)$. So once the front and back paths are simulated, the likelihood ratios can be calculated, stored, and reused for different derivatives. Also, compared to the LSMC (Longstaff and Schwartz, 2001) and to the kernel smoothing approach (Hong et al., 2017) for nested simulation, the GNS procedure does not require the selection of any basis function or any kernel function and the bandwidth parameter.

While reusing the same inner sample paths is expected to improve estimation efficiency and accuracy, it also introduces dependency among the loss estimators at all scenarios. The theoretical analysis for the GNS estimators (2.4) and (2.5) are not trivial, as shown in the next two sections.

References

Athanassios N Avramidis and Paul Hyden. Efficiency improvements for pricing american options with a stochastic mesh. In *Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 1*, pages 344–350. ACM, 1999.

Athanassios N Avramidis and Heinrich Matzinger. Convergence of the stochastic mesh estimator for pricing bermudan options. *Journal of Computational Finance*, 7(4):73–91, 2004.

Mark Broadie and Paul Glasserman. A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72, 2004.

Mark Broadie, Paul Glasserman, and Zachary Ha. Pricing american options by simulation using a stochastic mesh with optimized weights. In *Probabilistic Constrained Optimization*, pages 26–44. Springer, 2000.

Mark Broadie, Yiping Du, and Ciamac C Moallemi. Efficient risk estimation via nested sequential simulation. *Management Science*, 57(6):1172–1194, 2011.

Mark Broadie, Yiping Du, and Ciamac C Moallemi. Risk estimation via regression. *Operations Research*, 63(5):1077–1097, 2015.

Jacques F Carriere. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: mathematics and Economics*, 19(1):19–30, 1996.

Mingbin Feng and Jeremy Staum. Green simulation: Reusing the output of repeated experiments. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 27(4):23, 2017.

Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media, 2013.

Michael B Gordy and Sandeep Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.

L Jeff Hong, Sandeep Juneja, and Guangwu Liu. Kernel smoothing for nested estimation with application to portfolio risk measurement. *Operations Research*, 65(3):657–673, 2017.

Hai Lan, Barry L Nelson, and Jeremy Staum. A confidence interval procedure for expected shortfall risk measurement via two-level simulation. *Operations Research*, 58(5):1481–1490, 2010.

Shing-Hoi Lee. *Monte Carlo computation of conditional expectation quantiles.* PhD thesis, Stanford University, 1998.

Shing-Hoi Lee and Peter W Glynn. Computing the distribution function of a conditional expectation via monte carlo: Discrete conditioning spaces. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(3):238–258, 2003.

Guangwu Liu and L Jeff Hong. Revisit of stochastic mesh method for pricing american options. *Operations Research Letters*, 37(6):411–414, 2009.

Ming Liu and Jeremy Staum. Stochastic kriging for efficient nested simulation of expected shortfall. *Journal of Risk*, 12(3):3, 2010.

Francis A Longstaff and Eduardo S Schwartz. Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, 14(1):113–147, 2001.

Steven E Shreve. *Stochastic calculus for finance II: Continuous-time models*, volume 11. Springer Science & Business Media, 2004.

John N Tsitsiklis and Benjamin Van Roy. Regression methods for pricing complex american-style options. *IEEE Transactions on Neural Networks*, 12(4):694–703, 2001.

A Auxiliary proofs for results in Section ??

Proofs of Proposition **??**. Assumption **?? ??** assures that the likelihood ratio is well-defined so by Equation (2.3) we have

$$\mathbb{E}\left[L_m(\boldsymbol{s}_\tau)\right] = \mathbb{E}\left[\frac{1}{m}\sum_{j=1}^{m}\widehat{H}(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{S}}^{(j)}_{\tau^+})\right] = \mathbb{E}\left[\widehat{H}(\boldsymbol{s}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})\right] = L(\boldsymbol{s}_\tau).$$

Also, since $\mathbb{E}\left[|\widehat{H}|\right] < \infty$ and $\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}$, $j = 1, \ldots, m$ are i.i.d., by the strong law of large numbers we have $L_m(\boldsymbol{s}_\tau) \overset{a.s.}{\to} L(\boldsymbol{s}_\tau)$ as $m \to \infty$. This means that $\mathbb{P}\left(\lim_{m\to\infty} L_m(\boldsymbol{s}_\tau) = L(\boldsymbol{s}_\tau)\right) = 1$ for any fixed front $\boldsymbol{s}_\tau$. Since $\boldsymbol{S}_\tau$ and $\widetilde{\boldsymbol{S}}_{\tau^+}$ are independent by Assumption **?? ??**, so by the Independence Lemma (see Lemma 2.3.4 in Shreve, 2004, for example) we have

$$\begin{aligned}
\mathbb{P}\left(\lim_{m\to\infty} L_m(\boldsymbol{S}_\tau) = L(\boldsymbol{S}_\tau)\right) &= \mathbb{E}\left[\mathbb{1}\left\{\lim_{m\to\infty} L_m(\boldsymbol{S}_\tau) = L(\boldsymbol{S}_\tau)\right\}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}\left\{\lim_{m\to\infty} L_m(\boldsymbol{S}_\tau) = L(\boldsymbol{S}_\tau)\right\}|\boldsymbol{S}_\tau\right]\right] \\
&= \mathbb{E}\left[\mathbb{P}\left(\lim_{n\to\infty} L_m(\boldsymbol{s}_\tau) = L(\boldsymbol{s}_\tau)\right)\Big|_{\boldsymbol{s}_\tau = \boldsymbol{S}_\tau}\right] = 1.
\end{aligned}$$

This means that $L_m(\boldsymbol{S}_\tau) \overset{a.s.}{\to} L(\boldsymbol{S}_\tau)$ as $m \to \infty$ and the proof is complete. □

Proof of Lemma **??**. Note that

$$\begin{aligned}
\mathbb{E}\left[(R - \mathbb{E}\left[R|\mathcal{G}\right])^{2p}\right] \leq& \mathbb{E}\left[(|R| + |\mathbb{E}\left[R|\mathcal{G}\right]|)^{2p}\right] = \mathbb{E}\left[\sum_{k=0}^{2p}\binom{2p}{k}|R|^{2p-k}|\mathbb{E}\left[R|\mathcal{G}\right]|^k\right] \\
=& \mathbb{E}[R^{2p}] + \mathbb{E}\left[\mathbb{E}\left[R|\mathcal{G}\right]^{2p}\right] + \sum_{k=1}^{2p-1}\binom{2p}{k}\mathbb{E}\left[|R|^{2p-k}|\mathbb{E}\left[R|\mathcal{G}\right]|^k\right] \\
\overset{(*)}{\leq}& \mathbb{E}[R^{2p}] + \mathbb{E}\left[\mathbb{E}\left[R|\mathcal{G}\right]^{2p}\right] + \sum_{k=1}^{2p-1}\binom{2p}{k}(\mathbb{E}[R^{2p}])^{\frac{2p-k}{2p}}\left(\mathbb{E}\left[(\mathbb{E}\left[R|\mathcal{G}\right])^{2p}\right]\right)^{\frac{k}{2p}} \\
\overset{(**)}{\leq}& \mathbb{E}[R^{2p}] + \mathbb{E}[R^{2p}] + \sum_{k=1}^{2p-1}\binom{2p}{k}(\mathbb{E}\left[R^{2p}\right])^{\frac{2p-k}{2p}}(\mathbb{E}[R^{2p}])^{\frac{k}{2p}} \\
=& \sum_{k=0}^{2p}\binom{2p}{k}\mathbb{E}[R^{2p}] = 2^{2p}\mathbb{E}[R^{2p}],
\end{aligned}$$

where inequalities $(*)$ and $(**)$ follow from Hölder's and Jensen's inequalities, respectively. The proof is complete. □

Proof of Lemma **??**. According to the multinomial theorem and the conditional independence of R_j's, we have

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{1}{m}\sum_{j=1}^{m} R_j\right)^{2p}\right] &= \frac{1}{m^{2p}} \sum_{i_1+\cdots+i_k=2p} \frac{(2p)!}{i_1!i_2!\cdots i_k!} \mathbb{E}\left[R_{j_1}^{i_1}\cdots R_{j_k}^{i_k}\right] \\
&= \frac{1}{m^{2p}} \sum_{i_1+\cdots+i_k=2p} \frac{(2p)!}{i_1!i_2!\cdots i_k!} \mathbb{E}\left[\mathbb{E}\left[R_{j_1}^{i_1}\cdots R_{j_k}^{i_k}|\mathcal{G}\right]\right] \\
&= \frac{1}{m^{2p}} \sum_{i_1+\cdots+i_k=2p} \frac{(2p)!}{i_1!i_2!\cdots i_k!} \mathbb{E}\left[\mathbb{E}\left[R_{j_1}^{i_1}|\mathcal{G}\right]\cdots \mathbb{E}\left[R_{j_k}^{i_k}|\mathcal{G}\right]\right].
\end{aligned}$$

We will next bound the value and the number of summands. Since $i_1+\cdots+i_k=2p$, one can show that

$$\mathbb{E}\left[R_{j_1}^{i_1}\cdots R_{j_k}^{i_k}\right] \le \mathbb{E}\left[\left|R_{j_1}^{i_1}R_{j_2}^{i_2}\cdots R_{j_l}^{i_l}\right|\right] \overset{(*)}{\le} \left(\mathbb{E}\left[R_{j_1}^{2p}\right]\right)^{\frac{i_1}{2p}} \cdots \left(\mathbb{E}\left[R_{j_l}^{2p}\right]\right)^{\frac{i_l}{2p}} = \mathbb{E}\left[R_1^{2p}\right] < \infty,$$

where $(*)$ follows the generalized Hölder's inequality.

Since $\mathbb{E}\left[R_j|\mathcal{G}\right]=0$ for all $1\le j\le m$, for a summand to be non-zero it must have all $i_1,\ldots,i_k\ge 2$. Combine this with $i_1+\cdots+i_k=2p$, we have $k\le p$. Table 1 summarizes the multinomial coefficients and the number of summands of the form $\mathbb{E}\left[R_{j_1}^{i_1}\cdots R_{j_k}^{i_k}\right]$ for fixed numbers $k=1,\ldots,p$; the special case where $k=p$ is given in the second row.

summand expression	multinomial coefficient	# of different $\{i_1,\ldots,i_k\}$	# of different $\{j_1,\ldots,j_k\}$	product
$\mathbb{E}\left[R_{j_1}^{i_1}\cdots R_{j_k}^{i_k}\right]$	$\frac{(2p)!}{i_1!i_2!\cdots i_k!}$	# of integer solution satisfying $i_1,\ldots,i_k\ge 2$ and $i_1+\cdots+i_k=2p$. Does not depend on m.	$\binom{m}{k}=\mathcal{O}(m^k)$	$\mathcal{O}\left(m^k\right)\le\mathcal{O}(m^{p-1})$ for $k\le p-1$
$\mathbb{E}\left[R_{j_1}^{2}\cdots R_{j_p}^{2}\right]$	$\frac{(2p)!}{2^p}$	1	$\binom{m}{p}=\frac{m^p}{p!}+\mathcal{O}(m^{p-1})$	$c_pm^p+\mathcal{O}(m^{p-1})$ where $c_p=\frac{(2p)!}{2^p(p!)}$

Table 1: A breakdown of the number of summands for $k=1,\ldots,p$ unique of R_j's. The binomial coefficients are denoted by $\binom{n}{k}=\frac{n!}{(n-1)!k!}$.

For sufficiently large m, we have $\binom{m}{k} \leq \binom{m}{p}$ for $k \leq p$. Therefore, as $m \to \infty$,

$$\mathbb{E}\left[\left(\frac{1}{m}\sum_{j=1}^{m} R_j\right)^{2p}\right] = \frac{1}{m^{2p}}\left(c_p m^p + \mathcal{O}(m^{p-1})\right)\mathbb{E}\left[R_1^{2p}\right] = \mathcal{O}\left(m^{-p}\right).$$

The proof is complete. □

Proof of Theorem ??. Let $L_m(\boldsymbol{S}_\tau) - L(\boldsymbol{S}_\tau) = \frac{1}{m}\sum_{j=1}^{m} R_j$ where $R_j = H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(j)}) - L(\boldsymbol{S}_\tau)$ for $j = 1, ..., m$ and $\mathcal{G} = \sigma(\boldsymbol{S}_\tau)$ then it suffices to verify that the conditions of Lemma **??** hold.

Firstly, since $\widetilde{\boldsymbol{S}}_{\tau^+}^{(j)}$ are i.i.d. so R_j's are identically distributed and are conditional independent given $\boldsymbol{S}_\tau$. Moreover, by Equation (2.3) we have $\mathbb{E}\left[H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(j)})|\boldsymbol{S}_\tau\right] = L(\boldsymbol{S}_\tau)$ so $\mathbb{E}\left[R_j|\mathcal{G}\right] = 0$ for $j = 1, \ldots, m$. Lastly, the $2p$-moment of R_1 is bounded because

$$\mathbb{E}\left[|R_1|^{2p}\right] = \mathbb{E}\left[\left(H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(1)}) - \mathbb{E}\left[H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(1)})|\boldsymbol{S}_\tau\right]\right)^{2p}\right] \overset{(*)}{\leq} 4^p \mathbb{E}\left[\left|H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(1)})\right|^{2p}\right] < \infty,$$

where the inequality $(*)$ holds due to Lemma **??** with $R = H(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}^{(1)})$, and $\mathcal{G} = \sigma(\boldsymbol{S}_\tau)$. The proof is complete. □

B Auxiliary proofs for results in Sections ??–??

A few special instances of Cauchy-Schwartz's inequalities that are frequently used in our analysis so they are summarized in Lemma B.1 for ease of reference.

Lemma B.1. *For all vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of an inner product space, Cauchy-Schwartz's inequality asserts that $|\langle \boldsymbol{x}, \boldsymbol{y}\rangle|^2 \leq \langle \boldsymbol{x}, \boldsymbol{x}\rangle \cdot \langle \boldsymbol{y}, \boldsymbol{y}\rangle$, where $\langle\cdot,\cdot\rangle$ is the inner product. In particular, if $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y}$ is a vector of ones with compatible dimension, then*

$$\left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_{i=1}^{n} x_i^2. \tag{B.1}$$

Moreover, if $X_1, \ldots, X_n$ are identically distributed random variables, then

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] = \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \leq \frac{1}{n}\left(\sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]\right) = \mathbb{E}[X_1^2]. \tag{B.2}$$

Lastly, defining the inner product of two random variables, as the expectation of their product, then

$$\mathbb{E}\left[|XY|\right] \leq \left(\mathbb{E}\left[|X|^2\right]\right)^{1/2}\left(\mathbb{E}\left[|Y|^2\right]\right)^{1/2}. \tag{B.3}$$

Recall the smooth approximation function (**??**)

$$g_\epsilon(x) = \int_{-\infty}^{x/\epsilon} \phi(u)du,$$

where $\phi(u) = \frac{1}{4\pi}(1-\cos(u)) \cdot \mathbb{1}\{|u| \le 2\pi\}$. For any $\epsilon > 0$, $g_\epsilon(x)$ is bounded and differentiable with $g_\epsilon''(x) = \frac{1}{\epsilon^2}\phi'\left(\frac{x}{\epsilon}\right)$ and $\phi'(u) = \frac{1}{4\pi}\sin(u)\mathbb{1}\{|u| \le 2\pi\}$. Therefore,

$$\int_0^\infty \phi'(u)du = \int_0^{2\pi} \frac{1}{4\pi}\sin(u)du = 0, \text{ and} \tag{B.4a}$$

$$\int_0^\infty u|\phi'(u)|du = \int_0^{2\pi} \frac{1}{4\pi}u|\sin(u)|du = 1. \tag{B.4b}$$

Proposition **??**, Proposition **??**, and Theorem **??** are mostly analyzed in Sections **??**, **??**, and **??**, respectively. To complete the analysis therein, it remains to prove Lemma **??**.

Proof of Lemma **??**. For Equation (**??**), note that

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}\{L_m \ge 0\} - \mathbb{1}\{L \ge 0\}\right] &= \int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{\infty} p_m(y,z)dydz - \int_{\mathbb{R}}\int_0^\infty p_m(y,z)dydz \\
&= \int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{0} p_m(y,z)dydz \\
&\overset{(*)}{=} \int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{0} \left[p_m(0,z) + y \cdot \frac{\partial}{\partial y}p_m(u_y,z)\right]dydz \\
&= \int_{\mathbb{R}} \frac{z}{\sqrt{m}}p_m(0,z)dz + \int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{0} |y|\frac{\partial}{\partial y}p_m(u_y,z)dydz.
\end{aligned}$$

where $(*)$ holds by Assumption **??**. The first term on the RHS equals $\frac{\widetilde{p}(y)}{\sqrt{m}}\mathbb{E}[Z_m|L=0]$. This term equals zero because by Proposition **??**,

$$\mathbb{E}[Z_m|L=0] = \mathbb{E}[\mathbb{E}[L_m(\boldsymbol{S}_\tau) - L(\boldsymbol{S}_\tau)|L(\boldsymbol{S}_\tau) = 0, \boldsymbol{S}_\tau]] = 0.$$

The second term is of order $\mathcal{O}(m^{-1})$ because by Assumption **??**(D3), it is bounded by

$$\int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{0} |y| \cdot \bar{p}_{1,m}(z)dydz = \frac{1}{2m}\int_{\mathbb{R}} z^2\bar{p}_{1,m}(z)dz = \mathcal{O}(m^{-1}).$$

Equations (**??**) is equal to

$$\mathbb{E}\left[\mathbb{1}\{L_m \ge 0\} - \mathbb{1}\{L \ge 0\}\right] + 2\mathbb{E}\left[\mathbb{1}\{L \ge 0\} \cdot (\mathbb{1}\{L \ge 0\} - \mathbb{1}\{L_m \ge 0\})\right]$$

so it suffices to show that the second term is of order $\mathcal{O}(m^{-1})$. Since the indicator function and g_ϵ are both bounded using dominated convergence theorem we have

$$\mathbb{E}\left[\mathbb{1}\{L \ge 0\} \cdot (\mathbb{1}\{L \ge 0\} - \mathbb{1}\{L_m \ge 0\})\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\mathbb{1}\{L \geq 0\} \cdot \lim_{\epsilon \to 0} (g_\epsilon(L_m) - g_\epsilon(L))\right] \\
&= \lim_{\epsilon \to 0} \mathbb{E}\left[\mathbb{1}\{L \geq 0\} \cdot (g_\epsilon(L_m) - g_\epsilon(L))\right] \\
&= \lim_{\epsilon \to 0} \mathbb{E}\left[\mathbb{1}\{L \geq 0\} \cdot g_\epsilon'(L)(L - L_m)\right] + \lim_{\epsilon \to 0} \mathbb{E}\left[\mathbb{1}\{L \geq 0\} \cdot \frac{g_\epsilon''(L)}{2}(L - L_m)^2\right] \qquad \text{(B.5)} \\
&\quad + \lim_{\epsilon \to 0} \mathbb{E}\left[\text{remainder terms}\right]. \qquad \text{(B.6)}
\end{aligned}
$$

Here we omit the lengthy discussions on the technical assumptions needed to ensure that the expectation of the remainder term in (B.6) is of order $\mathcal{O}(m^{-3/2})$ and focus on the two expectations in (B.5). The first expectation equals zero by similar conditioning argument to (**??**). For the second expectation, note that

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{1}\{L \geq 0\} \cdot g_\epsilon''(L) \cdot (L - L_m)^2\right] \\
= &\int_{\mathbb{R}} \int_0^\infty \frac{1}{\epsilon^2} \phi'\left(\frac{y}{\epsilon}\right) \left(\frac{z}{\sqrt{m}}\right)^2 p_m(y, z) dy dz = \int_{\mathbb{R}} \int_0^\infty \frac{1}{\epsilon} \phi'(y) \left(\frac{z^2}{m}\right) p_m(\epsilon y, z) dy dz \\
\overset{(*)}{=} &\int_{\mathbb{R}} \int_0^\infty \frac{1}{\epsilon} \phi'(y) \left(\frac{z^2}{m}\right) \left[p_m(0, z) + \epsilon y \frac{\partial}{\partial y} p_m(y_\epsilon, z)\right] dy dz, \qquad \exists y_\epsilon \in [0, \epsilon y] \\
\leq &\int_{\mathbb{R}} \int_0^\infty \frac{1}{\epsilon} \phi'(y) \left(\frac{z^2}{m}\right) p_m(0, z) dy dz + \int_{\mathbb{R}} \int_0^\infty y |\phi'(y)| \left(\frac{z^2}{m}\right) \bar{p}_{1,m}(z) dy dz \\
= &\frac{1}{\epsilon m} \left(\int_0^{2\pi} \phi'(y)\, dy\right) \left(\int_{\mathbb{R}} z^2 p_m(0, z) dz\right) + \frac{1}{m} \left(\int_0^{2\pi} y |\phi'(y)| dy\right) \left(\int_{\mathbb{R}} z^2 \bar{p}_{1,m}(z) dz\right) \\
\overset{(**)}{\leq} &\, 0 + \mathcal{O}(m^{-1}),
\end{aligned}
$$

where $(*)$ and $(**)$ hold by Assumption **??** and Equations (B.4), respectively.

For Equation (**??**), note that

$$
\begin{aligned}
&\mathbb{E}\left[|L_m \cdot (\mathbb{1}\{L_m \geq 0\} - \mathbb{1}\{L \geq 0\})|\right] \\
= &\mathbb{E}\left[|(L + Z_m) \cdot (\mathbb{1}\{L_m \geq 0 > L\} - \mathbb{1}\{L \geq 0 > L_m\})|\right] \\
\leq &\mathbb{E}\left[|L + Z_m| \cdot \mathbb{1}\{L_m \geq 0 > L\}\right] + \mathbb{E}\left[|L + Z_m| \cdot \mathbb{1}\{L \geq 0 > L_m\}\right] \\
= &\left[\int_{-\infty}^0 \int_{z/\sqrt{m}}^0 \left|y + \frac{z}{\sqrt{m}}\right| p_m(y, z) dy dz\right] + \left[\int_0^\infty \int_0^{z/\sqrt{m}} \left|y + \frac{z}{\sqrt{m}}\right| p_m(y, z) dy dz\right] \\
\leq &\left[\int_{-\infty}^0 \int_{z/\sqrt{m}}^0 \left(|y| + \frac{|z|}{\sqrt{m}}\right) \bar{p}_{0,m}(z) dy dz\right] + \left[\int_0^\infty \int_0^{z/\sqrt{m}} \left(|y| + \frac{|z|}{\sqrt{m}}\right) \bar{p}_{0,m}(z) dy dz\right] \\
\leq &\left[\int_{-\infty}^0 \left(\frac{z^2}{2m} + \frac{z^2}{m}\right) \bar{p}_{0,m}(z) dz\right] + \left[\int_0^\infty \left(\frac{z^2}{2m} + \frac{z^2}{m}\right) \bar{p}_{0,m}(z) dz\right] \\
= &\frac{3}{2m} \int_{\mathbb{R}} z^2 \bar{p}_{0,m}(z) dz = \mathcal{O}(m^{-1}),
\end{aligned}
$$

$\rho_{mn} =$

where the last equality holds by Assumption **??**. The proof is complete. □

The variance bound (**??**) results from the following derivation.

$$
\begin{aligned}
\mathrm{Var}[\rho_{mn}] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} g\left(L_m^{(i)}\right) - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
=&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right) + \frac{1}{n}\sum_{i=1}^{n} g\left(L^{(i)}\right) - \mathbb{E}\left[g\left(L\right)\right] + \mathbb{E}\left[g\left(L\right)\right] - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
\overset{(*)}{\leq}& 3\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} g\left(L^{(i)}\right) - \mathbb{E}\left[g\left(L\right)\right]\right)^2 + \left(\mathbb{E}\left[g\left(L\right)\right] - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
\overset{(**)}{\leq}& 3\mathbb{E}\left[\left(g\left(L_m\right) - g\left(L\right)\right)^2\right] + \frac{3}{n}\mathbb{E}\left[\left(g\left(L\right) - \mathbb{E}\left[g\left(L\right)\right]\right)^2\right] + 3\mathbb{E}\left[\left(g\left(L\right) - g\left(L_m\right)\right)^2\right] \\
=&\frac{3}{n}\mathrm{Var}\left[g\left(L\right)\right] + 6\mathbb{E}\left[\left(g\left(L_m\right) - g\left(L\right)\right)^2\right] \qquad \text{(B.7)}
\end{aligned}
$$

where $(*)$ holds by inequality (B.1) and $(**)$ holds by applying inequality (B.2), Equation (**??**) in Lemma **??**, and Jensen's inequality to the three terms, respectively. Note that the terms in the second square on the RHS of $(*)$ are i.i.d. but those in the first square are not independent. So different inequalities are applied and the results have different orders.

C Auxiliary proofs for results in Section ??

The first part of Theorem **??**, i.e., the asymptotic normality of ρ_{mn}, is analyzed in Section **??**. We will use a few lemmas below to establish the convergences of the variance estimators.

Lemma C.1. *Suppose the conditions for Theorem **??** or those for Theorem **??** hold, then*

$$\mathbb{E}[|g(L_m) - g(L)|] = \mathcal{O}(m^{-1/2}), \text{ and} \qquad \text{(C.1)}$$

$$\mathbb{E}[|(g(L_m))^2 - (g(L))^2|] = \mathcal{O}(m^{-1/2}). \qquad \text{(C.2)}$$

Proof of Lemma C.1. For Equation (C.1), using Jensen's inequality we have that

$$\mathbb{E}[|g(L_m) - g(L)|] = \mathbb{E}[((g(L_m) - g(L))^2)^{1/2}] \leq \left(\mathbb{E}[(g(L_m) - g(L))^2]\right)^{-1/2} = \mathcal{O}(m^{-1/2}),$$

where the last equality holds due to Equation (**??**), (**??**), or (**??**).

For Equation (C.2), note that

$$\mathbb{E}[|(g(L_m))^2 - (g(L))^2|] = \mathbb{E}[|(g(L_m) - g(L))^2 + 2g(L)(g(L_m) - g(L))|]$$

$$\leq \mathbb{E}\left[(g(L_m)-g(L))^2\right] + 2\left(\mathbb{E}\left[(g(L))^2\right]\right)^{1/2}\left(\mathbb{E}\left[(g(L_m)-g(L))^2\right]\right)^{1/2}$$
$$= \mathcal{O}(m^{-1}) + \mathcal{O}(m^{-1/2}),$$

where the last equality holds due to Equation (??), (??), or (??). The proof is complete. □

Lemma C.2. *Suppose the conditions for Theorem* ?? *or those for Theorem* ?? *hold, then* $\widehat{\sigma}^2_{1,mn} \xrightarrow{p} \sigma^2_1$ *as* $\min\{m,n\} \to 0$.

Proof of Lemma C.2. In light of the Slutsky's theorem, it suffices to show that the two terms in $\widehat{\sigma}^2_{1,mn}$ converges in probability to the corresponding terms in σ^2_1.

For the first term in $\widehat{\sigma}^2_{1,mn}$, note that

$$\frac{1}{n}\sum_{i=1}^{n}(g(L_m^{(i)}))^2 - \mathbb{E}\left[(g\left(L\right))^2\right] = \frac{1}{n}\sum_{i=1}^{n}\left[(g(L_m^{(i)}))^2 - (g(L^{(i)}))^2\right] \tag{C.3}$$
$$+\left(\frac{1}{n}\sum_{i=1}^{n}(g(L^{(i)}))^2 - \mathbb{E}\left[(g\left(L\right))^2\right]\right). \tag{C.4}$$

Because $(g(L^{(i)}))^2$, $i = 1, \ldots, n$ are i.i.d. samples of $(g\left(L\right))^2$, the difference (C.4) converges to zero in probability by the weak law of large numbers. Also, by Lemma C.1 we have

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left[(g(L_m^{(i)}))^2 - (g(L^{(i)}))^2\right]\right|\right] \leq \mathbb{E}[|(g(L_m))^2 - (g(L))^2|] = \mathcal{O}(m^{-1/2}).$$

This means that (C.3) converges, in $\mathcal{L}^1$ and thus in probability, to zero as $m \to \infty$.

Consider the second term in $\widehat{\sigma}^2_{1,mn}$. By the continuous mapping theorem it suffices to show that $\frac{1}{n}\sum_{i=1}^{n} g(L_m^{(i)}) \xrightarrow{p} \mathbb{E}\left[g\left(L\right)\right]$. Replacing $(g(L_m^{(i)}))^2$, $(g(L^{(i)}))^2$, and $(g(L))^2$ in the previous steps by $g(L_m^{(i)})$, $g(L^{(i)})$, and $g(L)$ result in the desired conclusion. The proof is complete. □

For the convenience to state and prove the next lemma, we define the following notations:

$$R^{(j)} := \mathbb{E}[g'(L)\widehat{H}|\widetilde{\boldsymbol{S}}_{\tau^+} = \widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}],\ \widehat{R}^{(j)}_n := \frac{1}{n}\sum_{i=1}^{n} g'(L^{(i)})\widehat{H}^{(ij)},\ \text{and}\ \widehat{R}^{(j)}_{mn} := \frac{1}{n}\sum_{i=1}^{n} g'(L_m^{(i)})\widehat{H}^{(ij)}.$$

Lemma C.3. *Suppose the conditions for Theorem* ?? *hold, then the following equations hold.*

$$\lim_{m\to\infty}\mathbb{E}\left[\left|(\widehat{R}^{(1)}_{mn})^2 - (\widehat{R}^{(1)}_{n})^2\right|\right] = 0, \tag{C.5}$$
$$\lim_{n\to\infty}\mathbb{E}\left[\left|(\widehat{R}^{(1)}_{n})^2 - (R^{(1)})^2\right|\right] = 0,\ \textit{and} \tag{C.6}$$
$$\lim_{m\to\infty}\mathbb{E}\left[\left|g'(L_m)L_m - g'(L)L\right|\right] = 0. \tag{C.7}$$

Proof of Lemma C.3. Using inequality (B.2) we have that

$$\mathbb{E}[(\widehat{R}_{mn}^{(1)} - \widehat{R}_n^{(1)})^2] \leq \mathbb{E}[(g'(L_m)\widehat{H} - g'(L)\widehat{H})^2].$$

- For smooth functions, using Theorem **??** with $p = 2$ we have

$$\mathbb{E}[((g'(L_m) - g'(L))\widehat{H})^2] = \mathbb{E}[(g''(\Lambda_m)(L_m - L)\widehat{H})^2] \leq C_g^2(\mathbb{E}[(L_m - L)^4])^{1/2}(\mathbb{E}[\widehat{H}^4])^{1/2} = \mathcal{O}(m^{-1})$$

- For hockey-stick functions, $g'(x) = \mathbb{1}\{x \geq 0\} < \infty$, so by the dominated convergence theorem,

$$\lim_{m\to\infty} \mathbb{E}[((\mathbb{1}\{L_m \geq 0\} - \mathbb{1}\{L \geq 0\})\widehat{H})^2] = \mathbb{E}\left[\lim_{m\to\infty} ((\mathbb{1}\{L_m \geq 0\} - \mathbb{1}\{L \geq 0\})\widehat{H})^2\right] = 0,$$

  where the last equality holds because $L_m \overset{a.s}{\to} L$ as $m \to \infty$ according to Proposition **??**.

The moment conditions in Theorem **??** ensures that $\mathbb{E}[(g'(L)\widehat{H})^2] < \infty$. Also $\mathbb{E}\left[(\widehat{R}_n^{(1)})^2\right] \leq \mathbb{E}[(g'(L)\widehat{H})^2]$ by inequality (B.2). Therefore, for (C.5) we have

$$\begin{aligned}
&\lim_{m\to\infty} \mathbb{E}[|(\widehat{R}_{mn}^{(1)})^2 - (\widehat{R}_n^{(1)})^2|] \\
\leq &\lim_{m\to\infty} 2\mathbb{E}[(\widehat{R}_{mn}^{(1)} - \widehat{R}_n^{(1)})^2] + 2\left(\mathbb{E}\left[(\widehat{R}_n^{(1)})^2\right]\right)^{1/2}\left(\mathbb{E}[(\widehat{R}_{mn}^{(1)} - \widehat{R}_n^{(1)})^2]\right)^{1/2} = 0.
\end{aligned}$$

Given $\widetilde{\boldsymbol{S}}_{\tau^+}^{(1)}$, $g'(L^i)\widehat{H}^{(i1)}$ for $i = 1, \ldots, n$ are conditionally independent and identically distributed with mean $R^{(1)}$. Then, $\mathbb{E}[(\widehat{R}_n^{(1)} - R^{(1)})^2] = \mathcal{O}(n^{-1})$ by Lemma **??**. Finally, by Jensen's inequality we have $\mathbb{E}\left[(R^{(1)})^2\right] \leq \mathbb{E}[E[(g'(L)\widehat{H})^2|\widetilde{\boldsymbol{S}}_{\tau^+}]] = \mathbb{E}[(g'(L)\widehat{H})^2] < \infty$ so

$$\begin{aligned}
\mathbb{E}[|(\widehat{R}_n^{(1)})^2 - (R^{(1)})^2|] &\leq 2\mathbb{E}[(\widehat{R}_n^{(1)} - R^{(1)})^2] + 2\left(\mathbb{E}\left[(R^{(1)})^2\right]\right)^{1/2}\left(\mathbb{E}[(\widehat{R}_n^{(1)} - R^{(1)})^2]\right)^{1/2} \\
&= \mathcal{O}(n^{-1}) + \mathcal{O}(n^{-1/2}).
\end{aligned}$$

For (C.7), the term inside the absolute value equals $(g'(L_m) - g'(L))L_m - g'(L)(L - L_m)$, therefore

$$\mathbb{E}\left[\left|g'(L_m)L_m - g'(L)L\right|\right] \leq \mathbb{E}[|(g'(L_m) - g'(L))L_m|] + \mathbb{E}[|g'(L)(L - L_m)|] = \mathcal{O}(m^{-1/2}), \quad \text{(C.8)}$$

where the last equality holds because of the following:

- For smooth functions, the RHS of (C.8) equals

$$\begin{aligned}
&\mathbb{E}[|g''(\Lambda_m)(L_m - L)(L_m - L + L)|] + \mathbb{E}[|g'(L)(L - L_m)|] \\
\leq &C_g\left(\mathbb{E}[(L_m - L)^2] + \mathbb{E}[|(L_m - L)L|]\right) + \mathbb{E}[|g'(L)(L - L_m)|] \\
\leq &C_g\left(\mathbb{E}[(L_m - L)^2] + (\mathbb{E}[(L_m - L)^2])^{1/2}(\mathbb{E}[L^2])^{1/2}\right) + (\mathbb{E}[(g'(L))^2])^{1/2}(\mathbb{E}[(L_m - L)^2])^{1/2} \\
= &\mathcal{O}(m^{-1}) + \mathcal{O}(m^{-1/2}) + \mathcal{O}(m^{-1/2}).
\end{aligned}$$

- For the hockey-stick function, $g'(x) = \mathbb{1}\{x \geq 0\}$, so the RHS of (C.8) equals

$$
\begin{aligned}
&\mathbb{E}[|L_m \cdot (\mathbb{1}\{L_m \geq 0\} - \mathbb{1}\{L \geq 0\})|] + \mathbb{E}[|1\{L \geq 0\}(L_m - L)|] \\
\leq &\mathcal{O}(m^{-1}) + (\mathbb{E}[(L_m - L)])^{1/2} = \mathcal{O}(m^{-1}) + \mathcal{O}(m^{-1/2}).
\end{aligned}
$$

The proof is complete. □

Lemma C.4. *Suppose the conditions for Theorem* **??** *hold, then* $\widehat{\sigma}^2_{2,mn} \xrightarrow{p} \sigma_2^2$ *as* $\min\{m, n\} \to 0$.

Proof of Lemma C.4. In light of the Slutsky's theorem, it suffices to show that the two terms in $\widehat{\sigma}^2_{2,mn}$ converges in probability to the corresponding terms in σ_2^2.

For the first term in $\widehat{\sigma}^2_{2,mn}$, let $R = \mathbb{E}[g'(L)\widehat{H}|\widetilde{\boldsymbol{S}}_{\tau+}]$, then

$$
\begin{aligned}
\frac{1}{m}\sum_{j=1}^{m}(\widehat{R}^{(j)}_{mn})^2 - \mathbb{E}\left[R^2\right] = \frac{1}{m}\sum_{j=1}^{m}\left[(\widehat{R}^{(j)}_{mn})^2 - (\widehat{R}^{(j)}_{n})^2\right] + \frac{1}{m}\sum_{j=1}^{m}\left[(\widehat{R}^{(j)}_{n})^2 - (R^{(j)})^2\right] \quad &\text{(C.9)} \\
+ \left(\frac{1}{m}\sum_{j=1}^{m}(R^{(j)})^2 - \mathbb{E}\left[R^2\right]\right). \quad &\text{(C.10)}
\end{aligned}
$$

Because $(R^{(j)})^2$, $j = 1, \ldots, m$ are i.i.d. samples of R^2, the difference (C.10) converges to zero in probability by the weak law of large numbers. By Lemma C.3 we have

$$
\mathbb{E}\left[\left|\frac{1}{m}\sum_{j=1}^{m}\left[(\widehat{R}^{(j)}_{mn})^2 - (\widehat{R}^{(j)}_{n})^2\right]\right|\right] \leq \mathbb{E}[|(\widehat{R}^{(1)}_{mn})^2 - (\widehat{R}^{(1)}_{n})^2|] \to 0, \text{ as } m \to \infty, \text{ and}
$$

$$
\mathbb{E}\left[\left|\frac{1}{m}\sum_{j=1}^{m}\left[(\widehat{R}^{(j)}_{n})^2 - (R^{(j)})^2\right]\right|\right] \leq \mathbb{E}[|(\widehat{R}^{(1)}_{n})^2 - (R^{(1)})^2|] \to 0, \text{ as } n \to \infty.
$$

This means that both averages in (C.9) converges, in $\mathcal{L}^1$ and thus in probability, to zero as $\min\{m, n\} \to 0$.

Consider the second term in $\widehat{\sigma}^2_{2,mn}$. By the continuous mapping theorem it suffices to show that $\frac{1}{n}\sum_{i=1}^{n} g'(L^{(i)}_m)L^{(i)}_m \xrightarrow{p} \mathbb{E}\left[g'\left(L\right)L\right]$. Note that

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} g'(L^{(i)}_m)L^{(i)}_m - \mathbb{E}\left[g'\left(L\right)L\right] = \frac{1}{n}\sum_{i=1}^{n}\left[g'(L^{(i)}_m)L^{(i)}_m - g'(L^{(i)})L^{(i)}\right] \quad &\text{(C.11)} \\
+ \left(\frac{1}{n}\sum_{i=1}^{n} g'(L^{(i)})L^{(i)} - \mathbb{E}\left[g'\left(L\right)L\right]\right). \quad &\text{(C.12)}
\end{aligned}
$$

Because $g(L^{(i)})L^{(i)}$, $i = 1, \ldots, n$ are i.i.d. samples of $g(L)L$, the difference (C.12) converges to zero in probability by the weak law of large numbers. The $\mathcal{L}^1$-norm of (C.11) is bounded by $\mathbb{E}[|g'(L_m)L_m - g'(L)L|]$, which converges to zero by Lemma C.3. So (C.11) converges, in $\mathcal{L}^1$ and thus in probability, to zero as $m \to \infty$. The proof is complete. $\square$

D Auxiliary proofs for results in Section ??

For the indicator function $g(x) = \mathbb{1}\{x \geq 0\}$, we are able to establish analogous CLT results to Theorem **??** with an additional assumption. Assumption D.1 imposes regularity conditions so that limiting results as $\epsilon \to 0$ can be established. It is also a moment condition in disguise.

Assumption D.1. *The joint density $\psi(\boldsymbol{s}, y)$ of $(\boldsymbol{S}_\tau, L)$ exists. There exists a nonnegative function $\psi_1(\boldsymbol{s})$ such that $|\frac{\partial}{\partial y}\psi(\boldsymbol{s}, y)| \leq \psi_1(\boldsymbol{s})$ in an open neighborhood of $(\boldsymbol{s}, 0)$ for all $\boldsymbol{s}$. Denote $\psi_0(\boldsymbol{s}) = \psi(\boldsymbol{s}, 0)$. Moreover, suppose that $\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau} \left|\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+})\right| \psi_i(\boldsymbol{s})\mathrm{d}\boldsymbol{s}\right)^2\right]$, $\mathbb{E}\left[\int_{\mathbb{R}^\tau} \widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+})^2 \psi_i(\boldsymbol{s})\mathrm{d}\boldsymbol{s}\right]$, and $\int_{\mathbb{R}^\tau} \psi_i(\boldsymbol{s})\mathrm{d}\boldsymbol{s}$ are all finite for $i = 0, 1$.*

Assumption D.2. *The joint density $q_m(y_1, y_2)$ of $(L_m^{(1)}, L_m^{(2)})$ exists. There exists a nonnegative number C_q such that $|\partial_y q_m(y_1, y_2)| \leq C_q$ in an open neighborhood of $(0, 0)$.*

In the following, we analyze the asymptotic normality of ρ_{mn}. Recall that the definition of g_ϵ is (**??**). We let ϵ depend on m, denoted by ϵ_m, in the analysis of CLT results, and ϵ_m satisfies

$$\epsilon_m \to 0,\ m\epsilon_m^2 \to \infty \text{ and } m\epsilon_m^3 \to 0,\ \text{as } m \to \infty. \tag{D.1}$$

Note that the risk estimator ρ_{mn} can be decomposed as

$$\begin{aligned}
\rho_{mn} =& \frac{1}{n}\sum_{i=1}^n \left[g(L_m^{(i)}) - g_{\epsilon_m}(L_m^{(i)})\right] + \frac{1}{n}\sum_{i=1}^n \left[g_{\epsilon_m}(L_m^{(i)}) - g_{\epsilon_m}(L^{(i)}) + g(L^{(i)})\right] + \frac{1}{n}\sum_{i=1}^n \left[g_{\epsilon_m}(L^{(i)}) - g(L^{(i)})\right] \\
=& \underbrace{\frac{1}{n}\sum_{i=1}^n \left[g'_{\epsilon_m}(L^{(i)})(L_m^{(i)} - L^{(i)}) + g(L^{(i)})\right]}_{=:\mathcal{U}_{\epsilon_m,mn},\text{ will show } \mathcal{U}_{\epsilon_m,mn} \xrightarrow{d} \mathcal{N}(\rho,\sigma_{mn})} + \underbrace{\frac{1}{n}\sum_{i=1}^n \left[g''_{\epsilon_m}(L^{(i)})(L_m^{(i)} - L^{(i)})^2\right]}_{=:B,\text{ will show } \xrightarrow{\mathcal{L}^1} 0} + \underbrace{\text{remainders}}_{\text{discussions omitted}} \\
& + \underbrace{\frac{1}{n}\sum_{i=1}^n \left[g_{\epsilon_m}(L^{(i)}) - g(L^{(i)})\right]}_{=:C,\text{ will show } \xrightarrow{\mathcal{L}^1} 0} + \underbrace{\frac{1}{n}\sum_{i=1}^n \left[g(L_m^{(i)}) - g_{\epsilon_m}(L_m^{(i)})\right]}_{=:D,\text{ will show } \xrightarrow{\mathcal{L}^1} 0} \\
=& \mathcal{U}_{\epsilon_m,mn} + r^a_{\epsilon_m,mn} + r^b_{\epsilon_m,mn} + r^c_{\epsilon_m,mn} + \text{ remainders,}
\end{aligned}$$

where

$$\mathcal{U}_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^{n}\left[g(L^{(i)}) + g'_{\epsilon_m}(L^{(i)})(L_m^{(i)} - L^{(i)})\right], \tag{D.2}$$

$$r^a_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^{n} g''_{\epsilon_m}(L^{(i)})(L_m^{(i)} - L^{(i)})^2, \tag{D.3}$$

$$r^b_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^{n}\left[g_{\epsilon_m}(L^{(i)}) - g(L^{(i)})\right], \tag{D.4}$$

$$r^c_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^{n}\left[g(L_m^{(i)}) - g_{\epsilon_m}(L_m^{(i)})\right]. \tag{D.5}$$

Note that $\mathcal{U}_{\epsilon_m,mn} + r^a_{\epsilon_m,mn} +$ remainders $= \frac{1}{n}\sum_{i=1}^{n}[g(L) + g_{\epsilon_m}(L_m^{(i)}) - g_{\epsilon_m}(L^{(i)})]$ before the Taylor expansion for g_{ϵ_m} is applied. In fact, $r_{\epsilon_m,mn}$ in (**??**) is $r^a_{\epsilon_m,mn} + r^b_{\epsilon_m,mn} + r^c_{\epsilon_m,mn} +$ remainders. Here we omit the lengthy discussions on the technical assumptions needed to ensure that the remainder term in (B.6) is negligible and focus on analyzing the other terms. With the following Lemmas D.1, D.2, D.3 and D.4, we show that $\sigma_{mn}^{-1}(\mathcal{U}_{\epsilon_m,mn} - \rho) \xrightarrow{d} \mathcal{N}(0,1)$ and all $r^a_{\epsilon_m,mn}$, $r^b_{\epsilon_m,mn}$ and $r^c_{\epsilon_m,mn}$ converges to zero quickly. In Lemmas D.5 and D.6, we show that $\widehat{\sigma}^2_{1,mn}$ and $\widehat{\widetilde{\sigma}}^2_{2,mn}$ converge to σ_1^2 and $\widetilde{\sigma}_2^2$, respectively. Based these lemmas, the proof of Theorem **??** is complete.

Lemma D.1. *Suppose the conditions for Theorem* **??** *hold. Then,*

$$\sigma_{mn}^{-1}(\mathcal{U}_{\epsilon_m,mn} - \rho) \xrightarrow{d} \mathcal{N}(0,1), \text{ as } \min\{m,n\} \to \infty,$$

where

$$\sigma^2_{mn} = \frac{\sigma_1^2}{n} + \frac{\widetilde{\sigma}_2^2}{m} \tag{D.6}$$

*with $\widetilde{\sigma}_2^2$ defined by (**??**).*

Proof of Lemma D.1. Note that $\mathcal{U}_{\epsilon_m,mn}$ can be written as $\mathcal{U}_{\epsilon_m,mn} = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}(\Lambda_{\epsilon_m}^{(ij)} + g^{(i)})$, where

$$\Lambda_{\epsilon_m}^{(ij)} = g'_{\epsilon_m}(L^{(i)})(\widehat{H}^{(ij)} - L^{(i)}), \text{ and } g^{(i)} = g(L^{(i)}).$$

For notational convenience, we write simply Λ_{ϵ_m} in places for $\Lambda_{\epsilon_m}^{(11)}$. Furthermore,

$$\mathbb{E}[\Lambda_{\epsilon_m}|\boldsymbol{S}_\tau] = g'_{\epsilon_m}(L(\boldsymbol{S}_\tau))\left(\mathbb{E}[\widehat{H}(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+})|\boldsymbol{S}_\tau] - L(\boldsymbol{S}_\tau)\right) = 0. \tag{D.7}$$

Because the mapping $\Lambda_{\epsilon_m}^{(ij)}$ depends on ϵ_m, $\mathcal{U}_{\epsilon_m,mn}$ is not a U-statistic in Definition **??**. Nonetheless, we will show that it has similar asymptotic properties as a U-statistic, i.e., Lemma **??**.

Evidently, we know that $\mathbb{E}[\Lambda_n^{(ij)}] = 0$ and $\mathbb{E}[g^{(i)}] = \rho$. So $\mathbb{E}[\mathcal{U}_{\epsilon_m,mn}] = \rho$, i.e., U_n is an unbiased estimator of ρ.

Consider a related random variable (Hoeffdings decomposition)

$$\widetilde{\mathcal{U}}_{\epsilon_m,mn} = \sum_{i=1}^{n} \mathbb{E}[\mathcal{U}_{\epsilon_m,mn} - \rho | \boldsymbol{S}_\tau^{(i)}] + \sum_{j=1}^{m} \mathbb{E}[\mathcal{U}_{\epsilon_m,mn} - \rho | \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}] =: \widetilde{\mathcal{U}}_{\epsilon_m,n} + \widetilde{\mathcal{U}}_{\epsilon_m,m}.$$

By (D.7), it is easy to check that

$$\widetilde{\mathcal{U}}_{\epsilon_m,n} = \frac{1}{n}\sum_{i=1}^{n} g^{(i)} - \rho, \quad \widetilde{\mathcal{U}}_{\epsilon_m,m} = \frac{1}{m}\sum_{j=1}^{m} \mathbb{E}[\Lambda_{\epsilon_m}^{(1j)} | \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}],$$

and $\widetilde{\mathcal{U}}_{\epsilon_m,n}$ and $\widetilde{\mathcal{U}}_{\epsilon_m,m}$ are independent random variables.

By the classic CLT, clearly

$$\sqrt{n}\widetilde{\mathcal{U}}_{\epsilon_m,n} \xrightarrow{d} \mathcal{N}(0, \sigma_1^2), \tag{D.8}$$

as $n \to \infty$, where $\sigma_1^2 = \mathrm{Var}[g(L)]$.

For convenience, define i.i.d. random variables $Y_{\epsilon_m,j} = \mathbb{E}[\Lambda_{\epsilon_m}^{(1j)} | \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}]$ for $j = 1, \ldots, m$, and denote $Y_{\epsilon_m} = \mathbb{E}[\Lambda_{\epsilon_m} | \widetilde{\boldsymbol{S}}_{\tau+}]$. One can show that $\mathbb{E}[Y_{\epsilon_m,1}] = 0$. Furthermore, note that by Assumption D.1,

$$\begin{aligned}
Y_{\epsilon_m}^2 &= (\mathbb{E}[g'_{\epsilon_m}(L)(\widehat{H} - L) | \widetilde{\boldsymbol{S}}_{\tau+}])^2 \\
&= \left(\int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \frac{1}{\epsilon_m}\phi(\frac{y}{\epsilon_m})(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+}) - y)\psi(\boldsymbol{s}, y)\mathrm{d}\boldsymbol{s}\mathrm{d}y\right)^2 \\
&= \left(\int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \phi(u)(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+}) - \epsilon_m u)\psi(\boldsymbol{s}, \epsilon_m u)\mathrm{d}\boldsymbol{s}\mathrm{d}u\right)^2.
\end{aligned}$$

By Taylor expansion,

$$\psi(\boldsymbol{s}, \epsilon_m u) = \psi(\boldsymbol{s}, 0) + \epsilon_m u \frac{\partial}{\partial y}\psi(\boldsymbol{s}, \bar{u}) \tag{D.9}$$

where $\bar{u}$ is between 0 and $\epsilon_m u$, it follows that

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \phi(u)(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+}) - \epsilon_m u)\psi(\boldsymbol{s}, \epsilon_m u)\mathrm{d}\boldsymbol{s}\mathrm{d}u \\
=& \int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \phi(u)(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+}) - \epsilon_m u)\left[\psi(\boldsymbol{s}, 0) + \epsilon_m u \frac{\partial}{\partial y}\psi(\boldsymbol{s}, \bar{u})\right]\mathrm{d}\boldsymbol{s}\mathrm{d}u \\
=& \int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \phi(u)\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+})\psi(\boldsymbol{s}, 0)\mathrm{d}\boldsymbol{s}\mathrm{d}u + \epsilon_m \int_{\mathbb{R}}\int_{\mathbb{R}} \phi(u)u\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+})\frac{\partial}{\partial y}\psi(\boldsymbol{s}, \bar{u})\mathrm{d}\boldsymbol{s}\mathrm{d}u \\
&- \epsilon_m \int_{\mathbb{R}}\int_{\mathbb{R}^\tau} \phi(u)u\left[\psi(\boldsymbol{s}, 0) + \epsilon_m u \frac{\partial}{\partial y}\psi(\boldsymbol{s}, \bar{u})\right]\mathrm{d}\boldsymbol{s}\mathrm{d}u.
\end{aligned}$$

Due to some simple algebra and the finiteness of $\int_{\mathbb{R}^\tau}\psi(\boldsymbol{s},0)\mathrm{d}\boldsymbol{s}$, $\int_{\mathbb{R}^\tau}\psi_1(\boldsymbol{s})\mathrm{d}\boldsymbol{s}$, $\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau}\left|\widehat{H}(\boldsymbol{s},\widetilde{\boldsymbol{S}}_{\tau^+})\right|\psi_1(\boldsymbol{s})\mathrm{d}\boldsymbol{s}\right)^2\right]$ and $\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau}\left|\widehat{H}(\boldsymbol{s},\widetilde{\boldsymbol{S}}_{\tau^+})\right|\psi(\boldsymbol{s},0)\mathrm{d}\boldsymbol{s}\right)^2\right]$ in Assumption D.1, we obtain that

$$\mathbb{E}[Y_{\epsilon_m}^2]=\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau}\widehat{H}(\boldsymbol{s},\widetilde{\boldsymbol{S}}_{\tau^+})\psi(\boldsymbol{s},0)\mathrm{d}\boldsymbol{s}\right)^2\right]+\mathcal{O}(\epsilon_m)=:\sigma_2^2+\mathcal{O}(\epsilon_m). \tag{D.10}$$

Since $Y_{\epsilon_m,j}$, $j=1,\dots,m$ are i.i.d. samples, the characteristic function for $\sqrt{m}\widetilde{\mathcal{U}}_{\epsilon_m,m}$ is given by

$$\varphi_{\epsilon_m,m}(t)=\mathbb{E}\left[\exp\left\{it\sum_{j=1}^m\frac{Y_{\epsilon_m,j}}{\sqrt{m}}\right\}\right]=\left(\mathbb{E}\left[\exp\left\{it\frac{Y_{\epsilon_m,j}}{\sqrt{m}}\right\}\right]\right)^m.$$

Note that, by Taylor's theorem and $\mathbb{E}Y_{\epsilon_m}=0$,

$$\begin{aligned}\mathbb{E}\left[\exp\left\{it\frac{Y_{\epsilon_m,j}}{\sqrt{m}}\right\}\right]=&1-\frac{t^2}{2m}\mathbb{E}Y_{\epsilon_m}^2+o\left(\frac{t^2}{m}\right)\\=&1-\frac{t^2}{2m}\sigma_2^2+\frac{t^2}{2m}O(\epsilon_m)+o\left(\frac{t^2}{m}\right),\ \ \frac{t}{\sqrt{m}}\to 0,\ m,n\to\infty.\end{aligned}$$

This means that $\varphi_{\epsilon_m,m}(t)\to\exp\{-\frac{t^2}{2}\sigma_2^2\}$ as $m\to\infty$, so

$$\sqrt{m}\widetilde{\mathcal{U}}_{\epsilon_m,m}\xrightarrow{d}\mathcal{N}(0,\sigma_2^2) \tag{D.11}$$

by the Lévy's continuity theorem.

Since $\widetilde{\mathcal{U}}_{\epsilon_m,mn}=\widetilde{\mathcal{U}}_{\epsilon_m,n}+\widetilde{\mathcal{U}}_{\epsilon_m,m}$, $\widetilde{\mathcal{U}}_{\epsilon_m,n}$ and $\widetilde{\mathcal{U}}_{\epsilon_m,m}$ are independent, (D.8) and (D.11), it follows that as $\min\{m,n\}\to\infty$,

$$\frac{\widetilde{\mathcal{U}}_{\epsilon_m,mn}}{\sigma_{mn}}\xrightarrow{d}\mathcal{N}(0,1),\quad \sigma_{mn}^2=\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}, \tag{D.12}$$

where

$$\sigma_1^2=\mathrm{Var}[g(L)],\quad \sigma_2^2=\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau}\widehat{H}(\boldsymbol{s},\widetilde{\boldsymbol{S}}_{\tau^+})\psi(\boldsymbol{s},0)\mathrm{d}\boldsymbol{s}\right)^2\right]. \tag{D.13}$$

Next, note that

$$\frac{\mathcal{U}_{\epsilon_m,mn}-\rho}{\sigma_{mn}}=\frac{\widetilde{\mathcal{U}}_{\epsilon_m,mn}}{\sigma_{mn}}+\frac{\mathcal{U}_{\epsilon_m,mn}-\rho-\widetilde{\mathcal{U}}_{\epsilon_m,mn}}{\sigma_{mn}}. \tag{D.14}$$

It remains to show that $\sigma_{mn}^{-1}(\mathcal{U}_{\epsilon_m,mn}-\rho-\widetilde{\mathcal{U}}_{\epsilon_m,mn})\xrightarrow{d}0$ and it suffices to show that

$$\mathbb{E}\left[\sigma_{mn}^{-2}(\mathcal{U}_{\epsilon_m,mn}-\rho-\widetilde{\mathcal{U}}_{\epsilon_m,mn})^2\right]\to 0,\ \text{as}\ \min\{m,n\}\to\infty.$$

Note that

$$\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn}-\rho-\widetilde{\mathcal{U}}_{\epsilon_m,mn})^2\right]=\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn}-\rho)^2\right]+\mathbb{E}\left[\widetilde{\mathcal{U}}_{\epsilon_m,mn}^2\right]-2\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn}-\rho)\widetilde{\mathcal{U}}_{\epsilon_m,mn}\right]. \tag{D.15}$$

We investigate the three terms on the RHS of (D.15) via (i), (ii) and (iii) in the following, respectively.

(i) It follows from the definition of $\mathcal{U}_{\epsilon_m,mn}$ that

$$\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn}-\rho)^2\right]=\mathbb{E}\left[\left(\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\left\{\Lambda_{\epsilon_m}^{(ij)}+g^{(i)}\right\}-\rho\right)^2\right]$$

$$=\mathbb{E}\left[\left(\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\right)^2\right]+\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)^2\right]+2\mathbb{E}\left[\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)\right].$$

We examine these three terms on the RHS of this equation one by one. Let $\mathcal{G}=\sigma\left\{\boldsymbol{S}_\tau^{(1)},...,\boldsymbol{S}_\tau^{(n)}\right\}$.

First, it follows that

$$\mathbb{E}\left[\left(\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\right)^2\right]=\mathbb{E}\left\{\mathbb{E}\left[\left|\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\right|^2\middle|\mathcal{G}\right]\right\}=\mathbb{E}\left\{\frac{1}{m}\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\Lambda_{\epsilon_m}^{(i1)}\right|^2\middle|\mathcal{G}\right]\right\}$$

$$=\frac{1}{m}\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\Lambda_{\epsilon_m}^{(i1)}\right)^2\right]=\frac{1}{mn^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\Lambda_{\epsilon_m}^{(i1)}\right)^2\right]+\frac{1}{mn^2}\sum_{i\neq k}\mathbb{E}\left[\Lambda_{\epsilon_m}^{(i1)}\Lambda_{\epsilon_m}^{(k1)}\right]$$

$$=\frac{1}{mn}\mathbb{E}\left[\left(\Lambda_{\epsilon_m}^{(11)}\right)^2\right]+\frac{n-1}{mn}\mathbb{E}\left[\Lambda_{\epsilon_m}^{(11)}\Lambda_{\epsilon_m}^{(21)}\right]=\frac{1}{mn}\mathbb{E}\left[\left(\Lambda_{\epsilon_m}^{(11)}\right)^2\right]+\frac{n-1}{mn}\mathbb{E}\left\{\mathbb{E}\left[\Lambda_{\epsilon_m}^{(11)}\Lambda_{\epsilon_m}^{(21)}\middle|\widetilde{\boldsymbol{S}}_{\tau^+}^{(1)}\right]\right\}$$

$$=\frac{1}{mn}\mathbb{E}\left[\left(\Lambda_{\epsilon_m}^{(11)}\right)^2\right]+\frac{n-1}{mn}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda_{\epsilon_m}^{(11)}\middle|\widetilde{\boldsymbol{S}}_{\tau^+}^{(1)}\right]\right)^2\right],$$

where the second equality is by Lemma **??** because given $\mathcal{G}$, $\frac{1}{n}\sum_{i=1}^{n}\Lambda_{\epsilon_m}^{(ij)}$ $(j=1,...,m)$ are i.i.d samples with mean 0.

Second, by Lemma **??**,

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)^2\right]=\frac{1}{n}\mathbb{E}\left[\left(g^{(1)}-\rho\right)^2\right].$$

Third, due to (D.7), we have

$$\mathbb{E}\left[\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)\right]=\mathbb{E}\left\{\mathbb{E}\left[\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)\middle|\mathcal{G}\right]\right\}$$

$$=\mathbb{E}\left\{\mathbb{E}\left[\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\Lambda_{\epsilon_m}^{(ij)}\middle|\mathcal{G}\right]\left(\frac{1}{n}\sum_{i=1}^{n}g^{(i)}-\rho\right)\right\}=0.$$

Finally, we obtian

$$\mathbb{E}\left[\left(\widetilde{\mathcal{U}}_{\epsilon_m,mn}-\rho\right)^2\right]=\frac{1}{mn}\mathbb{E}\left[\left(\Lambda_{\epsilon_m}^{(11)}\right)^2\right]+\frac{n-1}{mn}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda_{\epsilon_m}^{(11)}\middle|\widetilde{\boldsymbol{S}}_{\tau^+}^{(1)}\right]\right)^2\right]+\frac{1}{n}\mathbb{E}\left[\left(g^{(1)}-\rho\right)^2\right].\tag{D.16}$$

(ii) Because $\widetilde{\mathcal{U}}_{\epsilon_m,mn} = \widetilde{\mathcal{U}}_{\epsilon_m,n} + \widetilde{\mathcal{U}}_{\epsilon_m,m}$, $\widetilde{\mathcal{U}}_{\epsilon_m,n}$ and $\widetilde{\mathcal{U}}_{\epsilon_m,m}$ are independent, and $\mathbb{E}\left[\widetilde{\mathcal{U}}_{\epsilon_m,n}\right] = \mathbb{E}\left[\widetilde{\mathcal{U}}_{\epsilon_m,m}\right] = 0$, we have

$$\begin{aligned}\mathbb{E}\left[\widetilde{\mathcal{U}}^2_{\epsilon_m,mn}\right] =&\mathbb{E}\left(\widetilde{\mathcal{U}}_{\epsilon_m,n} + \widetilde{\mathcal{U}}_{\epsilon_m,m}\right)^2 = \mathbb{E}\left[\widetilde{\mathcal{U}}^2_{\epsilon_m,n}\right] + \mathbb{E}\left[\widetilde{\mathcal{U}}^2_{\epsilon_m,m}\right] + 2\mathbb{E}\left(\widetilde{\mathcal{U}}_{\epsilon_m,n}\widetilde{\mathcal{U}}_{\epsilon_m,m}\right)\\ =&\mathbb{E}\left[\widetilde{\mathcal{U}}^2_{\epsilon_m,n}\right] + \mathbb{E}\left[\widetilde{\mathcal{U}}^2_{\epsilon_m,m}\right] = \frac{1}{n}\mathbb{E}\left(g^{(1)} - \rho\right)^2 + \frac{1}{m}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda^{(11)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(1)}_{\tau^+}\right]\right)^2\right].\end{aligned} \tag{D.17}$$

(iii) Note that

$$\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,mn}\right] = \mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,n}\right] + \mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,m}\right].$$

For the two terms on the RHS of above equation, first,

$$\begin{aligned}&\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,n}\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left\{(\mathcal{U}_{\epsilon_m,mn} - \rho)\left(g^{(i)} - \rho\right)\right\}\\ =&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left\{\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\left(g^{(i)} - \rho\right)\middle|\boldsymbol{S}^{(i)}_{\tau}\right]\right\} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left\{\left(g^{(i)} - \rho\right)\mathbb{E}\left[\mathcal{U}_{\epsilon_m,mn} - \rho\middle|\boldsymbol{S}^{(i)}_{\tau}\right]\right\}\\ =&\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left\{\left(g^{(i)} - \rho\right)\cdot\frac{1}{n}\left(g^{(i)} - \rho\right)\right\} = \frac{1}{n}\mathbb{E}\left[\left(g^{(1)} - \rho\right)^2\right].\end{aligned}$$

Second,

$$\begin{aligned}&\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,m}\right] = \frac{1}{m}\sum_{j=1}^m \mathbb{E}\left\{(\mathcal{U}_{\epsilon_m,mn} - \rho)\mathbb{E}\left[\Lambda^{(1j)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\right\}\\ =&\frac{1}{m}\sum_{j=1}^m \mathbb{E}\left\{\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\mathbb{E}\left[\Lambda^{(1j)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\right\}\\ =&\frac{1}{m}\sum_{j=1}^m \mathbb{E}\left\{\mathbb{E}\left[\Lambda^{(1j)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\mathbb{E}\left[\mathcal{U}_{\epsilon_m,mn} - \rho\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\right\}\\ =&\frac{1}{m}\sum_{j=1}^m \mathbb{E}\left\{\mathbb{E}\left[\Lambda^{(1j)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\cdot\frac{1}{m}\mathbb{E}\left[\Lambda^{(1j)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(j)}_{\tau^+}\right]\right\} = \frac{1}{m}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda^{(11)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(1)}_{\tau^+}\right]\right)^2\right].\end{aligned}$$

Therefore,

$$\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho)\widetilde{\mathcal{U}}_{\epsilon_m,mn}\right] = \frac{1}{n}\mathbb{E}\left(g^{(1)} - \rho\right)^2 + \frac{1}{m}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda^{(11)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(1)}_{\tau^+}\right]\right)^2\right]. \tag{D.18}$$

Finally, combining (D.15), (D.16), (D.17) and (D.18), we obtain that

$$\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho - \widetilde{\mathcal{U}}_{\epsilon_m,mn})^2\right] = \frac{1}{mn}\mathbb{E}\left[\Lambda^2_{\epsilon_m}\right] - \frac{1}{mn}\mathbb{E}\left[\left(\mathbb{E}\left[\Lambda^{(11)}_{\epsilon_m}\middle|\widetilde{\boldsymbol{S}}^{(1)}_{\tau^+}\right]\right)^2\right].$$

Note that

$$
\begin{aligned}
\mathbb{E}[\Lambda_{\epsilon_m}^2] &= \mathbb{E}\left[\mathbb{E}\left[\left(g_{\epsilon_m}'(L)(\widehat{H}(\boldsymbol{S}_\tau, \widetilde{\boldsymbol{S}}_{\tau^+}) - L)\right)^2 \middle| \widetilde{\boldsymbol{S}}_{\tau^+}\right]\right] \\
&= \mathbb{E}\left[\int_{\mathbb{R}}\int_{\mathbb{R}^\tau}\left(\frac{1}{\epsilon_m}\phi(\frac{y}{\epsilon_m})(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+}) - y)\right)^2 \psi(\boldsymbol{s}, y)\mathrm{d}\boldsymbol{s}\mathrm{d}y\right] \\
&= \epsilon_m^{-1}\mathbb{E}\left[\int_{\mathbb{R}}\int_{\mathbb{R}^\tau}\left(\phi(u)(\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+}) - \epsilon_m u)\right)^2 \psi(\boldsymbol{s}, \epsilon_m u)\mathrm{d}\boldsymbol{s}\mathrm{d}u\right].
\end{aligned}
$$

By the Taylor expansion (D.9), some simple algebra and the finiteness of $\mathbb{E}\left[\int_{\mathbb{R}^\tau}\widehat{H}^2(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+})\psi_1(\boldsymbol{s})\mathrm{d}\boldsymbol{s}\right]$, $\mathbb{E}\left[\int_{\mathbb{R}^\tau}\widehat{H}^2(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+})\psi(\boldsymbol{s}, 0)\mathrm{d}\boldsymbol{s}\right]$, $\mathbb{E}\left[\left(\int_{\mathbb{R}^\tau}\left|\widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau^+})\right|\psi(\boldsymbol{s}, 0)\mathrm{d}\boldsymbol{s}\right)^2\right]$, $\int_{\mathbb{R}^\tau}\psi(\boldsymbol{s}, 0)\mathrm{d}\boldsymbol{s}$ and $\int_{\mathbb{R}^\tau}\psi_1(\boldsymbol{s})\mathrm{d}\boldsymbol{s}$ in Assumption D.1, we obtain that

$$
\mathbb{E}[\Lambda_{\epsilon_m}^2] = \mathcal{O}(\epsilon_m^{-1}). \tag{D.19}
$$

By the definition of Y_{ϵ_m}, (D.10) and (D.19), it follows that

$$
\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho - \widetilde{\mathcal{U}}_{\epsilon_m,mn})^2\right] = \mathcal{O}\left(\frac{1}{mn\epsilon_m}\right) - \frac{\sigma_2^2 + \mathcal{O}(\epsilon_m)}{mn},
$$

and thus

$$
\sigma_{mn}^{-2}\mathbb{E}\left[(\mathcal{U}_{\epsilon_m,mn} - \rho - \widetilde{\mathcal{U}}_{\epsilon_m,mn})^2\right] = \sigma_{mn}^{-2}\left[\mathcal{O}\left(\frac{1}{mn\epsilon_m}\right) - \frac{\sigma_2^2 + \mathcal{O}(\epsilon_m)}{mn}\right] \to 0, \text{ as } m \to \infty,
$$

where the convergence is due to (D.1) and then $m\epsilon_m \to \infty$.

In summary, applying the Slutsky's theorem we have

$$
\frac{\mathcal{U}_{\epsilon_m,mn} - \rho}{\sigma_{mn}} \xrightarrow{d} \mathcal{N}(0, 1).
$$

The proof is complete. □

Lemma D.2. *Suppose the conditions for Theorem* **??** *hold. Then,*

$$
\sigma_{mn}^{-1}r_{\epsilon_m,mn}^a \xrightarrow{\mathcal{L}^1} 0, \text{ as } \min\{m, n\} \to \infty,
$$

where σ_{mn} is defined in (D.6).

Proof of Lemma D.2. Recall that $r_{\epsilon_m,mn}^a = \frac{1}{n}\sum_{i=1}^n\left[g_{\epsilon_m}''(L^{(i)})(L_m^{(i)} - L^{(i)})^2\right]$. By Assumption **??**,

$$
\mathbb{E}[|r_{\epsilon_m,mn}^a|] \leq \mathbb{E}\left[\left|g_{\epsilon_m}''(L)(L - L_m)^2\right|\right] = \int_{\mathbb{R}}\int_{-2\pi\epsilon_m}^{2\pi\epsilon_m}\left|\frac{1}{4\pi\epsilon_m^2}\sin\left(\frac{y}{\epsilon_m}\right)\left(\frac{z}{\sqrt{m}}\right)^2\right|p_m(y, z)\mathrm{d}y\mathrm{d}z
$$

$$= \frac{1}{4\pi m\epsilon_m}\int_{\mathbb{R}}\int_{-2\pi}^{2\pi} |\sin u|\, z^2 p_m(\epsilon_m u, z)\mathrm{d}u\mathrm{d}z \le \frac{1}{4\pi m\epsilon_m}\int_{-2\pi}^{2\pi}|\sin u|\,\mathrm{d}u\int_{\mathbb{R}} z^2\bar{p}_{0,m}(z)\mathrm{d}z = \mathcal{O}\left((m\epsilon_m)^{-1}\right).$$

Then as $m \to \infty$,

$$\mathbb{E}\left[\left|\sigma_{mn}^{-1} r^a_{\epsilon_m,mn}\right|\right] = \mathcal{O}\left(\left[\frac{m^2\epsilon_m^2}{n}\sigma_1^2 + m\epsilon_m^2\sigma_2^2\right]^{-1/2}\right) \to 0,$$

by (D.1). The proof is complete. □

Lemma D.3. *Suppose the conditions for Theorem* **??** *hold. Then,*

$$\sigma_{mn}^{-1} r^b_{\epsilon_m,mn} \xrightarrow{\mathcal{L}^1} 0, \text{ as } \min\{m,n\} \to \infty,$$

where σ_{mn} is defined in (D.6).

Proof of Lemma D.3. Recall that $r^b_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^n \left[g_{\epsilon_m}(L^{(i)}) - g(L^{(i)})\right]$. By the definition of $g_{\epsilon_m}(x)$, we know that

$$g_{\epsilon_m}(x) - g(x) = -\frac{1}{2}\left[\mathbb{1}\{0 \le x \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le x < 0\}\right] + \frac{1}{4\pi}\left[\frac{x}{\epsilon_m} - \sin\left(\frac{x}{\epsilon_m}\right)\right]\mathbb{1}\{|x| \le 2\pi\epsilon_m\}.$$

Therefore, it follows that

$$\begin{aligned} r^b_{\epsilon_m,mn} &= -\frac{1}{2}\frac{1}{n}\sum_{i=1}^n\left[\mathbb{1}\{0 \le L^{(i)} \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le L^{(i)} < 0\}\right] \\ &\quad + \frac{1}{4\pi}\frac{1}{n}\sum_{i=1}^n\left[\frac{L^{(i)}}{\epsilon_m} - \sin\left(\frac{L^{(i)}}{\epsilon_m}\right)\right]\mathbb{1}\{|L^{(i)}| \le 2\pi\epsilon_m\} \\ &= -\frac{1}{2}\frac{1}{n}\sum_{i=1}^n \bar{C}_i + \frac{1}{4\pi}\frac{1}{n}\sum_{i=1}^n \widetilde{C}_i, \end{aligned}$$

where

$$\bar{C}_i = \mathbb{1}\{0 \le L^{(i)} \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le L^{(i)} < 0\}, \text{ and } \widetilde{C}_i = \left[\frac{L^{(i)}}{\epsilon_m} - \sin\left(\frac{L^{(i)}}{\epsilon_m}\right)\right]\mathbb{1}\{|L^{(i)}| \le 2\pi\epsilon_m\}.$$

By Chebyshev's inequality, it suffices to prove that $\sigma_{mn}^{-1}\cdot\mathbb{E}[\bar{C}_1] \to 0$, $\sigma_{mn}^{-1}\cdot\mathbb{E}[\widetilde{C}_1] \to 0$, $\sigma_{mn}^{-2}n^{-1}\cdot\mathrm{Var}(\bar{C}_1) \to 0$ and $\sigma_{mn}^{-2}n^{-1}\cdot\mathrm{Var}(\widetilde{C}_1) \to 0$ as $\min\{m,n\} \to \infty$. Because by (D.1),

$$\sigma_{mn}^{-1}\epsilon_m^2 = \left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)^{-1/2}\cdot\epsilon_m^2 = \left(\frac{\sigma_1^2}{n\epsilon_m^4} + \frac{\sigma_2^2}{m\epsilon_m^4}\right)^{-1/2} \le \left(\frac{m\epsilon_m^4}{\sigma_2^2}\right)^{1/2} \to 0, \text{as } m \to \infty, \quad \text{(D.20)}$$

$$\sigma_{mn}^{-2}n^{-1} = \left(\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)^{-1}\cdot n^{-1} = \left(\sigma_1^2 + \frac{n\sigma_2^2}{m}\right)^{-1} < \sigma_1^{-2}, \quad \text{(D.21)}$$

we just have to prove that $\mathbb{E}[\bar{C}_1] = O(\epsilon_m^2)$, $\mathbb{E}[\widetilde{C}_1] = O(\epsilon_m^2)$, $\mathrm{Var}(\bar{C}_1) \to 0$ and $\mathrm{Var}(\widetilde{C}_1) \to 0$ as $\min\{m, n\} \to \infty$.

D.3.1. Proving $\mathbb{E}[\bar{C}_1] = O(\epsilon_m^2)$. Note that, by Assumption **??** and the mean value theorem,

$$\begin{aligned}\mathbb{E}[\bar{C}_1] =& \mathbb{E}\left[\mathbb{1}\{0 \le L \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le L < 0\}\right] = \mathbb{P}\left(0 \le L \le 2\pi\epsilon_m\right) - \mathbb{P}\left(-2\pi\epsilon_m \le L < 0\right)\\ =& \int_0^{2\pi\epsilon_m}\int_{\mathbb{R}} p_m(y,z)\mathrm{d}z\mathrm{d}y - \int_{-2\pi\epsilon_m}^{0}\int_{\mathbb{R}} p_m(y,z)\mathrm{d}z\mathrm{d}y\\ =& 2\pi\epsilon_m\int_{\mathbb{R}} p_m(\xi_1,z)\mathrm{d}z - 2\pi\epsilon_m\int_{\mathbb{R}} p_m(\xi_2,z)\mathrm{d}z = 2\pi\epsilon_m\int_{\mathbb{R}}\left[p_m(\xi_1,z) - p_m(\xi_2,z)\right]\mathrm{d}z\\ =& 2\pi\epsilon_m\int_{\mathbb{R}}(\xi_1-\xi_2)\frac{\partial}{\partial y}p_m(\xi,z)\mathrm{d}z,\end{aligned}$$

where $\xi_1 \in (0, 2\pi\epsilon_m)$, $\xi_2 \in (-2\pi\epsilon_m, 0)$, $\xi \in (\xi_2, \xi_1)$ and then $|\xi_1 - \xi_2| < 4\pi\epsilon_m$. Therefore,

$$\left|\mathbb{E}[\bar{C}_1]\right| \le 2\pi\epsilon_m\int_{\mathbb{R}}\left|(\xi_1-\xi_2)\frac{\partial}{\partial y}p_m(\xi,z)\right|\mathrm{d}z \le 8\pi\epsilon_m^2\left|\xi_1-\xi_2\right|\int_{\mathbb{R}}\bar{p}_{1,m}(z)\mathrm{d}z = \mathcal{O}(\epsilon_m^2).$$

D.3.2. Proving $\mathbb{E}[\widetilde{C}_1] = O(\epsilon_m^2)$. Note that, by integration by parts and the mean value theorem,

$$\begin{aligned}\mathbb{E}[\widetilde{C}_1] =& \mathbb{E}\left\{\left[\frac{L}{\epsilon_m} - \sin\left(\frac{L}{\epsilon_m}\right)\right]\mathbb{1}\{|L| \le 2\pi\epsilon_m\}\right\} = \int_{\mathbb{R}}\int_{-2\pi\epsilon_m}^{2\pi\epsilon_m}\left[\frac{y}{\epsilon_m} - \sin\left(\frac{y}{\epsilon_m}\right)\right]p_m(y,z)\mathrm{d}t\mathrm{d}z\\ =& \epsilon_m\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}(t-\sin t)p_m(\epsilon_m t,z)\mathrm{d}t\mathrm{d}z = \epsilon_m\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}(t-\sin t)\left[p_m(0,z) + \epsilon_m t\frac{\partial}{\partial y}p_m(\xi,z)\right]\mathrm{d}t\mathrm{d}z\\ =& \epsilon_m\int_{\mathbb{R}}p_m(0,z)\mathrm{d}z\int_{-2\pi}^{2\pi}(t-\sin t)\mathrm{d}t + \epsilon_m^2\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}(t-\sin t)t\frac{\partial}{\partial y}p_m(\xi,z)\mathrm{d}t\mathrm{d}z\\ =& 0 + \epsilon_m^2\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}(t-\sin t)t\frac{\partial}{\partial y}p_m(\xi,z)\mathrm{d}t\mathrm{d}z,\end{aligned}$$

where ξ is between 0 and $\epsilon_m t$. Hence,

$$\left|\mathbb{E}[\widetilde{C}_1]\right| \le \epsilon_m^2\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\left|(t-\sin t)t\frac{\partial}{\partial y}p_m(\xi,z)\right|\mathrm{d}t\mathrm{d}z \le \epsilon_m^2\int_{-2\pi}^{2\pi}\left|(t-\sin t)t\right|\mathrm{d}t\int_{\mathbb{R}}\bar{p}_{1,m}(z)\mathrm{d}z = \mathcal{O}(\epsilon_m^2).$$

D.3.3. Proving $\mathrm{Var}(\bar{C}_1) \to 0$. Note that, by integration by parts, the mean value theorem and Assumption **??**,

$$\begin{aligned}\mathrm{Var}(\bar{C}_1) \le& \mathbb{E}\left[|\bar{C}_1|^2\right] = \mathbb{E}\left|\mathbb{1}\{0 \le L \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le L < 0\}\right|^2\\ =& \mathbb{E}\left(\mathbb{1}^2\{0 \le L \le 2\pi\epsilon_m\} + \mathbb{1}^2\{-2\pi\epsilon_m \le L < 0\}\right) = \mathbb{P}\left(0 \le L \le 2\pi\epsilon_m\right) + \mathbb{P}\left(-2\pi\epsilon_m \le L < 0\right)\\ =& \int_{\mathbb{R}}\int_0^{2\pi\epsilon_m}p_m(y,z)\mathrm{d}y\mathrm{d}z + \int_{\mathbb{R}}\int_{-2\pi\epsilon_m}^{0}p_m(y,z)\mathrm{d}y\mathrm{d}z = 2\pi\epsilon_m\int_{\mathbb{R}}p_m(\xi_1,z)\mathrm{d}z + 2\pi\epsilon_m\int_{\mathbb{R}}p_m(\xi_2,z)\mathrm{d}z\end{aligned}$$

$$=2\pi\epsilon_m \int_{\mathbb{R}} [p_m(\xi_1, z) + p_m(\xi_2, z)]\, \mathrm{d}z \leq 4\pi\epsilon_m \int_{\mathbb{R}} \bar{p}_{0,m}(z)\mathrm{d}z = \mathcal{O}(\epsilon_m) \to 0,$$

where $\xi_1 \in (0, 2\pi\epsilon_m)$ and $\xi_2 \in (-2\pi\epsilon_m, 0)$.

D.3.4. Proving $\mathrm{Var}(\widetilde{C}_1) \to 0$. Note that, due to Assumption **??**,

$$\begin{aligned}
\mathrm{Var}(\widetilde{C}_1) \leq & \mathbb{E}\left[\widetilde{C}_1^2\right] = \mathbb{E}\left|\left[\frac{L}{\epsilon_m} - \sin\left(\frac{L}{\epsilon_m}\right)\right] \mathbb{1}\{|L| \leq 2\pi\epsilon_m\}\right|^2 \\
= & \int_{\mathbb{R}} \int_{-2\pi\epsilon_m}^{2\pi\epsilon_m} \left[\frac{y}{\epsilon_m} - \sin\left(\frac{y}{\epsilon_m}\right)\right]^2 p_m(y, z)\mathrm{d}y\mathrm{d}z \\
= & \epsilon_m \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t)^2\, p_m(\epsilon_m t, z)\mathrm{d}t\mathrm{d}z \\
\leq & \epsilon_m \int_{-2\pi}^{2\pi} (t - \sin t)^2\, \mathrm{d}t \int_{\mathbb{R}} \bar{p}_{0,m}(z)\mathrm{d}z = \mathcal{O}(\epsilon_m) \to 0.
\end{aligned}$$

In summary, we complete the proof. □

Lemma D.4. *Suppose the conditions for Theorem* **??** *hold. Then,*

$$\sigma_{mn}^{-1} r^c_{\epsilon_m, mn} \xrightarrow{\mathcal{L}^1} 0, \text{ as } \min\{m, n\} \to \infty,$$

where σ_{mn} *is defined in (D.6).*

Proof of Lemma D.4. Recall that $r^c_{\epsilon_m,mn} := \frac{1}{n}\sum_{i=1}^n \left[g(L_m^{(i)}) - g_{\epsilon_m}(L_m^{(i)})\right]$. Note that

$$\begin{aligned}
r^c_{\epsilon_m,mn} = & \frac{1}{2n}\sum_{i=1}^n \left[\mathbb{1}\left\{0 \leq L_m^{(i)} \leq 2\pi\epsilon_m\right\} - \mathbb{1}\left\{-2\pi\epsilon_m \leq L_m^{(i)} < 0\right\}\right] \\
& - \frac{1}{4\pi n}\sum_{i=1}^n \left[\frac{L_m^{(i)}}{\epsilon_m} - \sin\left(\frac{L_m^{(i)}}{\epsilon_m}\right)\right] \mathbb{1}\left\{|L_m^{(i)}| \leq 2\pi\epsilon_m\right\} \\
= & \frac{1}{2n}\sum_{i=1}^n \bar{D}_i - \frac{1}{4\pi n}\sum_{i=1}^n \widetilde{D}_i,
\end{aligned}$$

where

$$\begin{aligned}
\bar{D}_i := & \mathbb{1}\left\{0 \leq L_m^{(i)} \leq 2\pi\epsilon_m\right\} - \mathbb{1}\left\{-2\pi\epsilon_m \leq L_m^{(i)} < 0\right\}, \text{ and} \\
\widetilde{D}_i := & \left[\frac{L_m^{(i)}}{\epsilon_m} - \sin\left(\frac{L_m^{(i)}}{\epsilon_m}\right)\right] \mathbb{1}\left\{|L_m^{(i)}| \leq 2\pi\epsilon_m\right\}.
\end{aligned}$$

Therefore, by Chebyshev's inequality, it suffices to prove that $\sigma_{mn}^{-1} \cdot \mathbb{E}[\bar{D}_1] \to 0$ and $\sigma_{mn}^{-1} \cdot \mathbb{E}[\widetilde{D}_1] \to 0$, and $\mathrm{Var}(\sigma_{mn}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \bar{D}_i) \to 0$ and $\mathrm{Var}(\sigma_{mn}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \widetilde{D}_i) \to 0$ as $\min\{m, n\} \to \infty$.

D.4.1. Proving $\sigma_{mn}^{-1} \cdot \mathbb{E}[\bar{D}_1] \to 0$. Note that, by the mean value theorem and Assumption **??**,

$$
\begin{aligned}
\mathbb{E}[\bar{D}_1] =& \mathbb{E}\left[\mathbb{1}\{0 \le L_m \le 2\pi\epsilon_m\} - \mathbb{1}\{-2\pi\epsilon_m \le L_m < 0\}\right] \\
=& \mathbb{P}\left(0 \le L + Z_m/\sqrt{m} \le 2\pi\epsilon_m\right) - \mathbb{P}\left(-2\pi\epsilon_m \le L + Z_m/\sqrt{m} < 0\right) \\
=& \int_{\mathbb{R}} \int_{-z/\sqrt{m}}^{2\pi\epsilon_m - z/\sqrt{m}} p_m(y,z)\mathrm{d}y\mathrm{d}z - \int_{\mathbb{R}} \int_{-2\pi\epsilon_m - z/\sqrt{m}}^{-z/\sqrt{m}} p_m(y,z)\mathrm{d}y\mathrm{d}z \\
=& \int_{\mathbb{R}} 2\pi\epsilon_m p_m(\xi_1, z)\mathrm{d}z - \int_{\mathbb{R}} 2\pi\epsilon_m p_m(\xi_2, z)\mathrm{d}z \\
=& 2\pi\epsilon_m \int_{\mathbb{R}} [p_m(\xi_1, z) - p_m(\xi_2, z)]\mathrm{d}z = 2\pi\epsilon_m \int_{\mathbb{R}} (\xi_1 - \xi_2)\frac{\partial}{\partial y} p_m(\xi, z)\mathrm{d}z,
\end{aligned}
$$

where $\xi_1 \in (-z/\sqrt{m}, 2\pi\epsilon_m - z/\sqrt{m})$, $\xi_2 \in (-2\pi\epsilon_m - z/\sqrt{m}, -z/\sqrt{m})$ and $\xi \in (\xi_2, \xi_1)$. Then $|\xi_1 - \xi_2| < 4\pi\epsilon_m$. Therefore, by Assumption **??** and (D.20),

$$
\left|\sigma_{mn}^{-1}\mathbb{E}[\bar{D}_1]\right| \le 2\pi\sigma_{mn}^{-1}\epsilon_m \int_{\mathbb{R}} \left|(\xi_1 - \xi_2)\frac{\partial}{\partial y}p_m(\xi, z)\right| \mathrm{d}z \le 8\pi\sigma_{mn}^{-1}\epsilon_m^2 \int_{\mathbb{R}} \bar{p}_{1,m}(z)\mathrm{d}z = \mathcal{O}\left(\sigma_{mn}^{-1}\epsilon_m^2\right) \to 0,
$$

as $m \to \infty$.

D.4.2. Proving $\sigma_{mn}^{-1} \cdot \mathbb{E}[\widetilde{D}_1] \to 0$. Note that

$$
\begin{aligned}
\mathbb{E}[\widetilde{D}_1] =& \mathbb{E}\left\{\left[\frac{L_m}{\epsilon_m} - \sin\left(\frac{L_m}{\epsilon_m}\right)\right] \mathbb{1}\{|L_m| \le 2\pi\epsilon_m\}\right\} \\
=& \mathbb{E}\left\{\left[\frac{L + Z_m/\sqrt{m}}{\epsilon_m} - \sin\left(\frac{L + Z_m/\sqrt{m}}{\epsilon_m}\right)\right] \mathbb{1}\left\{|L + Z_m/\sqrt{m}| \le 2\pi\epsilon_m\right\}\right\} \\
=& \int_{\mathbb{R}} \int_{-2\pi\epsilon_m - z/\sqrt{m}}^{2\pi\epsilon_m - z/\sqrt{m}} \left[\frac{y + z/\sqrt{m}}{\epsilon_m} - \sin\left(\frac{y + z/\sqrt{m}}{\epsilon_m}\right)\right] p_m(y,z)\mathrm{d}y\mathrm{d}z \\
=& \epsilon_m \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t) p_m(\epsilon_m t - z/\sqrt{m}, z)\mathrm{d}t\mathrm{d}z.
\end{aligned}
$$

By Taylor expansion, we have

$$
p_m(\epsilon_m t - z/\sqrt{m}, z) = p_m(-z/\sqrt{m}, z) + \epsilon_m t \frac{\partial}{\partial y} p_m(\bar{y}, z), \tag{D.22}
$$

where $\bar{y} \in (-z/\sqrt{m}, 2\pi\epsilon_m t - z/\sqrt{m})$. Then

$$
\begin{aligned}
\mathbb{E}[\widetilde{D}_1] =& \epsilon_m \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t)\left[p_m(-z/\sqrt{m}, z) + \epsilon_m t \frac{\partial}{\partial y} p_m(\bar{y}, z)\right] \mathrm{d}t\mathrm{d}z \\
=& \epsilon_m \int_{\mathbb{R}} p_m(-z/\sqrt{m}, z)\mathrm{d}z \int_{-2\pi}^{2\pi} (t - \sin t)\mathrm{d}t + \epsilon_m^2 \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t) t \frac{\partial}{\partial y} p_m(\bar{y}, z)\mathrm{d}t\mathrm{d}z \\
=& 0 + \epsilon_m^2 \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t) t \frac{\partial}{\partial y} p_m(\bar{y}, z)\mathrm{d}t\mathrm{d}z.
\end{aligned}
$$

Therefore, by Assumption **??** and (D.20),

$$\begin{aligned}\left|\sigma_{mn}^{-1}\mathbb{E}[\widetilde{D}_1]\right| \leq& \sigma_{mn}^{-1}\epsilon_m^2 \int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\left|(t-\sin t)t\frac{\partial}{\partial y}p_m(\bar{y},z)\right|\mathrm{d}t\mathrm{d}z\\ \leq& \sigma_{mn}^{-1}\epsilon_m^2\int_{-2\pi}^{2\pi}|(t-\sin t)t|\,\mathrm{d}t\int_{\mathbb{R}}\bar{p}_{1,m}(z)\mathrm{d}z = \mathcal{O}\left(\sigma_{mn}^{-1}\epsilon_m^2\right)\to 0,\end{aligned}$$

as $m\to\infty$.

D.4.3. Proving $\mathrm{Var}(\sigma_{mn}^{-1}\cdot\frac{1}{n}\sum_{i=1}^n \bar{D}_i)\to 0$. Note that

$$\mathrm{Var}\left(\sigma_{mn}^{-1}\cdot\frac{1}{n}\sum_{i=1}^n\bar{D}_i\right) = \sigma_{mn}^{-2}\cdot\frac{1}{n}\mathrm{Var}(\bar{D}_1)+\sigma_{mn}^{-2}\cdot\frac{n-1}{n}\mathrm{Cov}(\bar{D}_1,\bar{D}_2).$$

By (D.21), we just have to prove that $\mathrm{Var}(\bar{D}_1)\to 0$ and $\sigma_{mn}^{-2}\cdot\mathrm{Cov}(\bar{D}_1,\bar{D}_2)\to 0$.

First, note that by the mean value theorem,

$$\begin{aligned}\mathrm{Var}(\bar{D}_1)\leq&\mathbb{E}\left[|\bar{D}_1|^2\right] = \mathbb{E}\left|\mathbb{1}\{0\leq L_m\leq 2\pi\epsilon_m\}-\mathbb{1}\{-2\pi\epsilon_m\leq L_m<0\}\right|^2\\ =&\mathbb{E}\left(\mathbb{1}^2\{0\leq L_m\leq 2\pi\epsilon_m\}+\mathbb{1}^2\{-2\pi\epsilon_m\leq L_m<0\}\right)\\ =&\mathbb{P}\left(0\leq L+Z_m/\sqrt{m}\leq 2\pi\epsilon_m\right)+\mathbb{P}\left(-2\pi\epsilon_m\leq L+Z_m/\sqrt{m}<0\right)\\ =&\int_{\mathbb{R}}\int_{-z/\sqrt{m}}^{2\pi\epsilon_m-z/\sqrt{m}}p_m(y,z)\mathrm{d}y\mathrm{d}z+\int_{\mathbb{R}}\int_{-2\pi\epsilon_m-z/\sqrt{m}}^{-z/\sqrt{m}}p_m(y,z)\mathrm{d}y\mathrm{d}z\\ =&\int_{\mathbb{R}}2\pi\epsilon_m p_m(\xi_1,z)\mathrm{d}z+\int_{\mathbb{R}}2\pi\epsilon_m p_m(\xi_2,z)\mathrm{d}z,\end{aligned}$$

where $\xi_1\in(-z/\sqrt{m},2\pi\epsilon_m-z/\sqrt{m})$ and $\xi_2\in(-2\pi\epsilon_m-z/\sqrt{m},-z/\sqrt{m})$. Then by Assumption **??**,

$$\mathrm{Var}(\bar{D}_1)\leq\int_{\mathbb{R}}2\pi\epsilon_m\bar{p}_{0,m}(z)\mathrm{d}z+\int_{\mathbb{R}}2\pi\epsilon_m\bar{p}_{0,m}(z)\mathrm{d}z = 4\pi\epsilon_m\int_{\mathbb{R}}\bar{p}_{0,m}(z)\mathrm{d}z=\mathcal{O}(\epsilon_m)\to 0,$$

as $m\to\infty$.

Second, we study $\sigma_{mn}^{-2}\cdot\mathrm{Cov}(\bar{D}_1,\bar{D}_2)\to 0$. Note that

$$\sigma_{mn}^{-2}\mathrm{Cov}(\bar{D}_1,\bar{D}_2)=\sigma_{mn}^{-2}\mathbb{E}[\bar{D}_1\bar{D}_2]-(\sigma_{mn}^{-1}\mathbb{E}[\bar{D}_1])^2.$$

We have proved $\sigma_{mn}^{-1}\mathbb{E}[\bar{D}_1]\to 0$ as $m\to\infty$ in **D.4.1**, it suffices to prove that $\sigma_{mn}^{-2}\cdot\mathbb{E}[\bar{D}_1\bar{D}_2]\to 0$ in the following.

Note that

$$\begin{aligned}\mathbb{E}[\bar{D}_1\bar{D}_2]=&\mathbb{E}\left[\bar{D}_1\left(\mathbb{1}\left\{0\leq L_m^{(2)}\leq 2\pi\epsilon_m\right\}-\mathbb{1}\left\{-2\pi\epsilon_m\leq L_m^{(2)}<0\right\}\right)\right]\\ =&\mathbb{E}\left[\bar{D}_1\mathbb{1}\left\{0\leq L_m^{(2)}\leq 2\pi\epsilon_m\right\}\right]-\mathbb{E}\left[\bar{D}_1\mathbb{1}\left\{-2\pi\epsilon_m\leq L_m^{(2)}<0\right\}\right].\end{aligned}$$

Evidently, $\mathbb{E}\left[\bar{D}_1 \mathbb{1}\left\{0 \leq L_m^{(2)} \leq 2\pi\epsilon_m\right\}\right]$ and $\mathbb{E}\left[\bar{D}_1 \mathbb{1}\left\{-2\pi\epsilon_m \leq L_m^{(2)} < 0\right\}\right]$ have the same convergence rate, so we only analyze the first one. Specifically, by the mean value theorem and Assumption D.2,

$$
\begin{aligned}
& \mathbb{E}\left[\bar{D}_1 \mathbb{1}\left\{0 \leq L_m^{(2)} \leq 2\pi\epsilon_m\right\}\right] \\
=& \mathbb{E}\left[\left(\mathbb{1}\left\{0 \leq L_m^{(1)} \leq 2\pi\epsilon_m\right\} - \mathbb{1}\left\{-2\pi\epsilon_m < L_m^{(1)} < 0\right\}\right) \mathbb{1}\left\{0 \leq L_m^{(2)} \leq 2\pi\epsilon_m\right\}\right] \\
=& \mathbb{E}\left[\mathbb{1}\left\{0 \leq L_m^{(1)} \leq 2\pi\epsilon_m\right\}\mathbb{1}\left\{0 \leq L_m^{(2)} \leq 2\pi\epsilon_m\right\}\right] - \mathbb{E}\left[\mathbb{1}\left\{-2\pi\epsilon_m < L_m^{(1)} < 0\right\}\mathbb{1}\left\{0 \leq L_m^{(2)} \leq 2\pi\epsilon_m\right\}\right] \\
=& \int_0^{2\pi\epsilon_m}\int_0^{2\pi\epsilon_m} q_m(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 - \int_0^{2\pi\epsilon_m}\int_{-2\pi\epsilon_m}^0 q_m(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 \\
=& 2\pi\epsilon_m \int_0^{2\pi\epsilon_m} q_m(\xi_1, y_2)\mathrm{d}y_2 - 2\pi\epsilon_m \int_0^{2\pi\epsilon_m} q_m(\xi_2, y_2)\mathrm{d}y_2 = 2\pi\epsilon_m \int_0^{2\pi\epsilon_m} \left[q_m(\xi_1, y_2) - q_m(\xi_2, y_2)\right]\mathrm{d}y_2 \\
=& 2\pi\epsilon_m \int_0^{2\pi\epsilon_m} \frac{\partial}{\partial y_1} q_m(\xi, y_2)(\xi_1 - \xi_2)\mathrm{d}y_2 = (2\pi\epsilon_m)^2 \frac{\partial}{\partial y_1} q_m(\xi, \xi')(\xi_1 - \xi_2),
\end{aligned}
$$

where $\xi_1 \in (0, 2\pi\epsilon_m)$, $\xi_2 \in (-2\pi\epsilon_m, 0)$, $\xi \in (\xi_2, \xi_1)$, $\xi' \in (0, 2\pi\epsilon_m)$ and so $|\xi_1 - \xi_2| < 4\pi\epsilon_m$. Therefore, by Assumption D.2 and (D.1), it follows that $\left|\sigma_{mn}^{-2} \cdot \mathbb{E}[\bar{D}_1\bar{D}_2]\right| = \mathcal{O}\left(\sigma_{mn}^{-2} \cdot \epsilon_m^3\right) \to 0$ as $m \to \infty$.

D.4.4. Proving $\mathrm{Var}(\sigma_{mn}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \widetilde{D}_i) \to 0$. Similar to the analysis in **D.4.3**, by (D.21) and

$$
\mathrm{Var}\left(\sigma_{mn}^{-1} \cdot \frac{1}{n}\sum_{i=1}^n \widetilde{D}_i\right) = \sigma_{mn}^{-2} \cdot \frac{1}{n}\mathrm{Var}(\widetilde{D}_1) + \sigma_{mn}^{-2} \cdot \frac{n-1}{n}\mathrm{Cov}(\widetilde{D}_1, \widetilde{D}_2),
$$

we just have to prove that $\mathrm{Var}(\widetilde{D}_1) \to 0$ and $\sigma_{mn}^{-2} \cdot \mathrm{Cov}(\widetilde{D}_1, \widetilde{D}_2) \to 0$. Note that

$$
\begin{aligned}
& \mathrm{Var}(\widetilde{D}_1) \leq \mathbb{E}\left[\widetilde{D}_1^2\right] = \mathbb{E}\left\{\left[\frac{L_m}{\epsilon_m} - \sin\left(\frac{L_m}{\epsilon_m}\right)\right]^2 \mathbb{1}\{|L_m| \leq 2\pi\epsilon_m\}\right\} \\
=& \mathbb{E}\left\{\left[\frac{L + Z_m/\sqrt{m}}{\epsilon_m} - \sin\left(\frac{L + Z_m/\sqrt{m}}{\epsilon_m}\right)\right]^2 \mathbb{1}\left\{|L + Z_m/\sqrt{m}| \leq 2\pi\epsilon_m\right\}\right\} \\
=& \int_{\mathbb{R}}\int_{-2\pi\epsilon_m - z/\sqrt{m}}^{2\pi\epsilon_m - z/\sqrt{m}} \left[\frac{y + z/\sqrt{m}}{\epsilon_m} - \sin\left(\frac{y + z/\sqrt{m}}{\epsilon_m}\right)\right]^2 p_m(y, z)\mathrm{d}y\mathrm{d}z \\
=& \epsilon_m \int_{\mathbb{R}}\int_{-2\pi}^{2\pi} (t - \sin t)^2 p_m(\epsilon_m t - z/\sqrt{m}, z)\mathrm{d}t\mathrm{d}z.
\end{aligned}
$$

By Taylor expansion (D.22),

$$
\begin{aligned}
& \mathrm{Var}(\widetilde{D}_1) \leq \epsilon_m \int_{\mathbb{R}}\int_{-2\pi}^{2\pi} (t - \sin t)^2 \left[p_m(-z/\sqrt{m}, z) + \epsilon_m t \frac{\partial}{\partial y} p_m(\bar{y}, z)\right]\mathrm{d}t\mathrm{d}z \\
=& \epsilon_m \int_{\mathbb{R}} p_m(-z/\sqrt{m}, z)\mathrm{d}z \int_{-2\pi}^{2\pi} (t - \sin t)^2\mathrm{d}t + \epsilon_m^2 \int_{\mathbb{R}}\int_{-2\pi}^{2\pi} (t - \sin t)^2 t \frac{\partial}{\partial y} p_m(\bar{y}, z)\mathrm{d}t\mathrm{d}z.
\end{aligned}
$$

Furthermore, by Assumption **??**,

$$\mathrm{Var}(\widetilde{D}_1) \le \epsilon_m \int_{\mathbb{R}} \bar{p}_{0,m}(z)\mathrm{d}z \int_{-2\pi}^{2\pi} (t - \sin t)^2 \mathrm{d}t + \epsilon_m^2 \int_{\mathbb{R}} \int_{-2\pi}^{2\pi} (t - \sin t)^2 |t| \bar{p}_{1,m}(z)\mathrm{d}t\mathrm{d}z = \mathcal{O}(\epsilon_m).$$

Hence, $\mathrm{Var}(\widetilde{D}_1) \to 0$ as $m \to \infty$.

Second, we study $\sigma_{mn}^{-2} \cdot \mathrm{Cov}(\widetilde{D}_1, \widetilde{D}_2) \to 0$. Note that

$$\sigma_{mn}^{-2}\mathrm{Cov}(\widetilde{D}_1, \widetilde{D}_2) = \sigma_{mn}^{-2}\mathbb{E}[\widetilde{D}_1\widetilde{D}_2] - (\sigma_{mn}^{-1}\mathbb{E}[\widetilde{D}_1])^2.$$

We have proved that $\sigma_{mn}^{-1}\mathbb{E}[\widetilde{D}_1] \to 0$ as $m \to \infty$ in **D.4.2**, so it suffices to prove that $\sigma_{mn}^{-2}\cdot\mathbb{E}[\widetilde{D}_1\widetilde{D}_2] \to 0$ in the following.

By Assumption D.2,

$$\begin{aligned}
&\mathbb{E}[\widetilde{D}_1\widetilde{D}_2] = \mathbb{E}\left\{\left[\frac{L_m^{(1)}}{\epsilon_m} - \sin\left(\frac{L_m^{(1)}}{\epsilon_m}\right)\right] \mathbb{1}\left\{|L_m^{(1)}| \le 2\pi\epsilon_m\right\} \cdot \left[\frac{L_m^{(2)}}{\epsilon_m} - \sin\left(\frac{L_m^{(2)}}{\epsilon_m}\right)\right] \mathbb{1}\left\{|L_m^{(2)}| \le 2\pi\epsilon_m\right\}\right\} \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \left[\frac{y_1}{\epsilon_m} - \sin\left(\frac{y_1}{\epsilon_m}\right)\right] \mathbb{1}\{|y_1| \le 2\pi\epsilon_m\} \left[\frac{y_2}{\epsilon_m} - \sin\left(\frac{y_2}{\epsilon_m}\right)\right] \mathbb{1}\{|y_2| \le 2\pi\epsilon_m\} q_m(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 \\
&= \epsilon_m^2 \int_{\mathbb{R}}\int_{\mathbb{R}} (u - \sin u)\mathbb{1}\{|u| \le 2\pi\}(v - \sin v)\mathbb{1}\{|v| \le 2\pi\} q_m(\epsilon_m u, \epsilon_m v)\mathrm{d}u\mathrm{d}v \\
&= \epsilon_m^2 \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} (u - \sin u)(v - \sin v) q_m(\epsilon_m u, \epsilon_m v)\mathrm{d}u\mathrm{d}v.
\end{aligned}$$

From the Taylor expansion

$$q_m(\epsilon_m u, \epsilon_m v) = q_m(0, \epsilon_m v) + \epsilon_m u \frac{\partial}{\partial y_1} q_m(\bar{u}, \epsilon_m v),$$

where $\bar{u}$ is between 0 and $\epsilon_m u$, then

$$\begin{aligned}
\mathbb{E}[\widetilde{D}_1\widetilde{D}_2] =& \epsilon_m^2 \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} (u - \sin u)(v - \sin v)\left[q_m(0, \epsilon_m v) + \epsilon_m u \frac{\partial}{\partial y_1} q_m(\bar{u}, \epsilon_m v)\right]\mathrm{d}u\mathrm{d}v \\
=& \epsilon_m^2 \int_{-2\pi}^{2\pi} (v - \sin v) q_m(0, \epsilon_m v)\mathrm{d}v \int_{-2\pi}^{2\pi} (u - \sin u)\mathrm{d}u \\
&+ \epsilon_m^3 \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} (u - \sin u)(v - \sin v) u \frac{\partial}{\partial y_1} q_m(\bar{u}, \epsilon_m v)\mathrm{d}u\mathrm{d}v \\
=& 0 + \epsilon_m^3 \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} (u - \sin u)(v - \sin v) u \frac{\partial}{\partial y_1} q_m(\bar{u}, \epsilon_m v)\mathrm{d}u\mathrm{d}v.
\end{aligned}$$

Furthermore, by Assumption D.2,

$$\left|\mathbb{E}[\widetilde{D}_1\widetilde{D}_2]\right| \le \epsilon_m^3 \int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi} \left|(u - \sin u)(v - \sin v) u \frac{\partial}{\partial y_1} q_m(\bar{u}, \epsilon_m v)\right| \mathrm{d}u\mathrm{d}v$$

$$\leq \epsilon_m^3 C_q \int_{-2\pi}^{2\pi} |(u-\sin u)u| \, \mathrm{d}u \int_{-2\pi}^{2\pi} |v-\sin v| \, \mathrm{d}v = \epsilon_m^3 C_q \left(\frac{16}{3}\pi^2 - 4\pi\right) 4\pi^2 = \mathcal{O}(\epsilon_m^3),$$

Combine with (D.1), then it follows that $\sigma_{mn}^{-2} \cdot \mathbb{E}[\widetilde{D}_1 \widetilde{D}_2] = \mathcal{O}\left(\sigma_{mn}^{-2} \cdot \epsilon_m^3\right) \to 0$ as $m \to \infty$.

In summary, we complete the proof. □

Lemma D.5. *Suppose the conditions for Theorem* **??** *hold, then* $\widehat{\sigma}_{1,mn}^2 \xrightarrow{p} \sigma_1^2$ *as* $\min\{m,n\} \to 0$.

In fact, the proof of Lemma D.5 is contained in that of Lemma C.2, so we do not present it here again.

Lemma D.6. *Suppose the conditions for Theorem* **??** *hold, then* $\widehat{\widetilde{\sigma}}_{2,mn}^2 \xrightarrow{p} \widetilde{\sigma}_2^2$ *as* $\min\{m,n\} \to 0$*, where* $\widetilde{\sigma}_2^2$ *and* $\widehat{\widetilde{\sigma}}_{2,mn}^2$ *are defined in (*??*) and (*??*), respectively.*

Proof of Lemma D.6. For convenience, define shorthanded notations

$$R := \int_{\mathbb{R}^\tau} \widehat{H}(s, \widetilde{\boldsymbol{S}}_{\tau+}) \psi(\boldsymbol{s}, 0) \mathrm{d}\boldsymbol{s}, \; R^{(j)} := \int_{\mathbb{R}^\tau} \widehat{H}(\boldsymbol{s}, \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}) \psi(\boldsymbol{s}, 0) \mathrm{d}\boldsymbol{s}, \; R_\epsilon^{(j)} := \mathbb{E}[g_\epsilon'(L) \widehat{H}^{(1j)} | \widetilde{\boldsymbol{S}}_{\tau+}^{(j)}], \tag{D.23}$$

and

$$\widehat{R}_\epsilon^{(j)} = \frac{1}{n} \sum_{i=1}^n g_\epsilon'(L^{(i)}) \widehat{H}^{(ij)}, \; \widehat{R}_{\epsilon,m}^{(j)} = \frac{1}{n} \sum_{i=1}^n g_\epsilon'(L_m^{(i)}) \widehat{H}^{(ij)}. \tag{D.24}$$

The random variables are identically distributed for $j = 1, \ldots, m$ so the superscript may be omitted for $j = 1$. Then we have

$$\begin{aligned}
\widehat{\widetilde{\sigma}}_{2,mn}^2 - \widetilde{\sigma}_2^2 =& \frac{1}{m} \sum_{j=1}^m \left(\widehat{R}_{\epsilon,m}^{(j)}\right)^2 - \mathbb{E}[R^2] \\
=& \frac{1}{m} \sum_{j=1}^m \left[\left(\widehat{R}_{\epsilon,m}^{(j)}\right)^2 - \left(\widehat{R}_\epsilon^{(j)}\right)^2\right] + \frac{1}{m} \sum_{j=1}^m \left[\left(\widehat{R}_\epsilon^{(j)}\right)^2 - \left(R_\epsilon^{(j)}\right)^2\right] \\
&+ \frac{1}{m} \sum_{j=1}^m \left[\left(R_\epsilon^{(j)}\right)^2 - \left(R^{(j)}\right)^2\right] + \left[\frac{1}{m} \sum_{j=1}^m \left(R^{(j)}\right)^2 - \mathbb{E}[R^2]\right] \\
=& A_1 + A_2 + A_3 + A_4,
\end{aligned}$$

where

$$\begin{aligned}
A_1 :=& \frac{1}{m} \sum_{j=1}^m \left[\left(\widehat{R}_{\epsilon,m}^{(j)}\right)^2 - \left(\widehat{R}_\epsilon^{(j)}\right)^2\right], \; A_2 := \frac{1}{m} \sum_{j=1}^m \left[\left(\widehat{R}_\epsilon^{(j)}\right)^2 - \left(R_\epsilon^{(j)}\right)^2\right], \\
A_3 :=& \frac{1}{m} \sum_{j=1}^m \left[\left(R_\epsilon^{(j)}\right)^2 - \left(R^{(j)}\right)^2\right], \; A_4 := \left[\frac{1}{m} \sum_{j=1}^m \left(R^{(j)}\right)^2 - \mathbb{E}[R^2]\right].
\end{aligned}$$

In the following, we will prove the convergence of A_i $(i = 1, 2, 3, 4)$ to 0 in $\mathcal{L}^1$ one by one.

D.6.1. Proving $A_1 \to 0$ in $\mathcal{L}^1$. Note that by (B.3),

$$\mathbb{E}\left[|A_1|\right] \leq \mathbb{E}\left[\left|\left(\widehat{R}^{(1)}_{\epsilon,m}\right)^2 - \left(\widehat{R}^{(1)}_{\epsilon}\right)^2\right|\right] \leq \mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} - \widehat{R}^{(1)}_{\epsilon}\right|^2\right) \cdot \mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} + \widehat{R}^{(1)}_{\epsilon}\right|^2\right). \tag{D.25}$$

Then we study the two terms on RHS of (D.25). First, by the mean value theorem, $|g''_\epsilon(\cdot)| \leq \frac{1}{4\pi\epsilon^2}$ and Lemma **??**,

$$\begin{aligned}
&\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} - \widehat{R}^{(1)}_{\epsilon}\right|^2\right) \leq \mathbb{E}\left[\left|\left[g'_\epsilon(L_m) - g'_\epsilon(L)\right]\widehat{H}\right|^2\right] \leq \frac{1}{16\pi\epsilon^4}\mathbb{E}\left[\left|(L_m - L)\,\widehat{H}\right|^2\right] \\
\leq&\frac{1}{16\pi\epsilon^4}\left(\mathbb{E}\left[|L_m - L|^4\right]\right)^{1/2}\left(\mathbb{E}\left[\widehat{H}^4\right]\right)^{1/2} = \frac{1}{16\pi\epsilon^4}\left(\mathcal{O}\left(\frac{1}{m^2}\right)\right)^{1/2}\left(\mathbb{E}\left[\widehat{H}^4\right]\right)^{1/2} = \mathcal{O}\left(\frac{1}{m\epsilon^4}\right).
\end{aligned}$$

Second, note that by (B.1),

$$\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} + \widehat{R}^{(1)}_{\epsilon}\right|^2\right) \leq 2\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} - \widehat{R}^{(1)}_{\epsilon}\right|^2\right) + 8\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon}\right|^2\right), \tag{D.26}$$

and similar to (D.19), we have $\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon}\right|^2\right) = \mathcal{O}(\epsilon^{-1})$.

Therefore, based on (D.25), $\mathbb{E}[|A_1|] \leq \mathcal{O}\left(\frac{1}{\epsilon^{5/2}m^{1/2}}\right) \to 0$ due to $m\epsilon^5 \to \infty$ as $m \to \infty$.

D.6.2. Proving $A_2 \to 0$ in $\mathcal{L}^1$. Note that by (B.3),

$$\mathbb{E}[|A_2|] \leq \mathbb{E}\left[\left|\left(\widehat{R}^{(j)}_{\epsilon}\right)^2 - \left(R^{(j)}_{\epsilon}\right)^2\right|\right] \leq \left(\mathbb{E}\left[\left|\widehat{R}^{(j)}_{\epsilon} - R^{(j)}_{\epsilon}\right|^2\right]\right)^{\frac{1}{2}} \cdot \left(\mathbb{E}\left[\left|\widehat{R}^{(j)}_{\epsilon} + R^{(j)}_{\epsilon}\right|^2\right]\right)^{\frac{1}{2}}. \tag{D.27}$$

For the first term on the RHS of (D.27), it follows from Lemma **??** that

$$\mathbb{E}\left(\left|\widehat{R}^{(j)}_{\epsilon} - R^{(j)}_{\epsilon}\right|^2\right) \leq \frac{1}{n} \cdot \mathbb{E}\left|g'_\epsilon(L)\widehat{H}\right|^2 = \mathcal{O}\left(\frac{1}{n\epsilon}\right),$$

where the equality holds similarly to (D.19). For the second term on the RHS of (D.27), similar to (D.26) and (D.19),

$$\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} + \widehat{R}^{(1)}_{\epsilon}\right|^2\right) \leq 2\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} - \widehat{R}^{(1)}_{\epsilon}\right|^2\right) + 8\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon}\right|^2\right) \leq 2\mathbb{E}\left(\left|\widehat{R}^{(1)}_{\epsilon,m} - \widehat{R}^{(1)}_{\epsilon}\right|^2\right) + \mathcal{O}\left(\frac{1}{\epsilon}\right).$$

Therefore, based on (D.27), $\mathbb{E}[|A_2|] \leq O(\frac{1}{n^{1/2}\epsilon}) \to 0$. That is, A_2 converges in L^1 to 0.

D.6.3. Proving $A_3 \to 0$ and $A_4 \to 0$ in $\mathcal{L}^1$. Due to Lemma **??**, these two results are evident.

The proof is complete. $\square$

E Variance analysis for the case of indicator function

Assumption E.1. *(D1) The joint density* $q_m(y_1, y_2, z_1, z_2)$ *of* $(L^{(1)}, L^{(2)}, \widetilde{Z}_m^{(1)}, \widetilde{Z}_m^{(2)})$ *and its partial derivatives* $\frac{\partial}{\partial y_i} q_m(y_1, y_2, z_1, z_2)$ $(i = 1, 2)$ *exist for every* m *and for all* (y_1, y_2, z_1, z_2).

(D2) For every $m \geq 1$, *there exist nonnegative functions* $\bar{q}_{k,m}(z_1, z_2), (k = 0, 1)$ *such that for* $i = 1, 2$,

$$q_m(y_1, y_2, z_1, z_2) \leq \bar{q}_{0,m}(z_1, z_2) \text{ and } \left| \frac{\partial}{\partial y_i} q_m(y_1, y_2, z_1, z_2) \right| \leq \bar{q}_{1,m}(z_1, z_2), \quad \forall (y_1, y_2, z_1, z_2).$$

(D3) For $i = 1, 2, k = 0, 1$ *and* $0 \leq r \leq 2$

$$\sup_m \int_{\mathbb{R}} |z_i|^r \bar{q}_{k,m}(z_1, z_2) dz_1 dz_2 < \infty.$$

E.1 Variance analysis

$$\begin{aligned}
&\mathrm{Var}[\rho_{mn}] = \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} g\left(L_m^{(i)}\right) - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
=&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right) + \frac{1}{n}\sum_{i=1}^{n} g\left(L^{(i)}\right) - \mathbb{E}\left[g\left(L\right)\right] + \mathbb{E}\left[g\left(L\right)\right] - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
\leq& 3\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} g\left(L^{(i)}\right) - \mathbb{E}\left[g\left(L\right)\right]\right)^2 + \left(\mathbb{E}\left[g\left(L\right)\right] - \mathbb{E}\left[g\left(L_m\right)\right]\right)^2\right] \\
=&3\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)\right)^2\right] + \frac{3}{n}\mathbb{E}\left[\left(g\left(L\right) - \mathbb{E}\left[g\left(L\right)\right]\right)^2\right] + 3(\mathrm{Bias}[\rho_{mn}])^2, \qquad \text{(E.1)}
\end{aligned}$$

where the inequality holds by inequality (B.1). By Proposition **??**, $\mathrm{Bias}[\rho_{mn}] = \mathcal{O}(m^{-1})$, therefore we just have to analyze the first term on the RHS of (E.1). Note that

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)\right)^2\right] \\
=&\frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^{n}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)^2 + \sum_{i\neq k}\left(g\left(L_m^{(i)}\right) - g\left(L^{(i)}\right)\right)\left(g\left(L_m^{(k)}\right) - g\left(L^{(k)}\right)\right)\right] \\
=&\frac{1}{n}\mathbb{E}\left[\left(g\left(L_m\right) - g\left(L\right)\right)^2\right] + \frac{n-1}{n}\mathbb{E}\left[\left(g\left(L_m^{(1)}\right) - g\left(L^{(1)}\right)\right)\left(g\left(L_m^{(2)}\right) - g\left(L^{(2)}\right)\right)\right] \\
\leq&\frac{1}{n} + \frac{n-1}{n}\mathbb{E}\left[\left(g\left(L_m^{(1)}\right) - g\left(L^{(1)}\right)\right)\left(g\left(L_m^{(2)}\right) - g\left(L^{(2)}\right)\right)\right], \qquad \text{(E.2)}
\end{aligned}$$

where the inequality is due to $g(\cdot) = \mathbb{1}\{\cdot \geq 0\} \leq 1$. Hence, it suffices to analyze the second on the RHS of (E.2). It follows that

$$
\begin{aligned}
&\mathbb{E}\left[\left(g\left(L_m^{(1)}\right) - g\left(L^{(1)}\right)\right)\left(g\left(L_m^{(2)}\right) - g\left(L^{(2)}\right)\right)\right] \\
=&\mathbb{E}\left[\left(\mathbb{1}\{L_m^{(1)} \geq 0\} - \mathbb{1}\{L^{(1)} \geq 0\}\right)\left(\mathbb{1}\{L_m^{(2)} \geq 0\} - \mathbb{1}\{L^{(2)} \geq 0\}\right)\right] \\
=&\mathbb{E}\left[\left(\mathbb{1}\{L_m^{(1)} \geq 0 > L^{(1)}\} - \mathbb{1}\{L^{(1)} \geq 0 > L_m^{(1)}\}\right)\left(\mathbb{1}\{L_m^{(2)} \geq 0 > L^{(2)}\} - \mathbb{1}\{L^{(2)} \geq 0 > L_m^{(2)}\}\right)\right] \\
=&\mathbb{E}\Big[\mathbb{1}\{L_m^{(1)} \geq 0 > L^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\} - \mathbb{1}\{L^{(1)} \geq 0 > L_m^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\} \\
&-\mathbb{1}\{L_m^{(1)} \geq 0 > L^{(1)}, L^{(2)} \geq 0 > L_m^{(2)}\} + \mathbb{1}\{L^{(1)} \geq 0 > L_m^{(1)}, L^{(2)} \geq 0 > L_m^{(2)}\}\Big] \\
=&\mathbb{P}\left(L_m^{(1)} \geq 0 > L^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\right) - \mathbb{P}\left(L^{(1)} \geq 0 > L_m^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\right) \\
&- \mathbb{P}\left(L_m^{(1)} \geq 0 > L^{(1)}, L^{(2)} \geq 0 > L_m^{(2)}\right) + \mathbb{P}\left(L^{(1)} \geq 0 > L_m^{(1)}, L^{(2)} \geq 0 > L_m^{(2)}\right).
\end{aligned}
$$

We only have to examine the first term on the RHS of the above equation. Note that

$$
\begin{aligned}
&\mathbb{P}\left(L_m^{(1)} \geq 0 > L^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\right) \\
=&\mathbb{P}\left(L^{(1)} + \widetilde{Z}_m^{(1)}/\sqrt{m} \geq 0 > L^{(1)}, L^{(2)} + \widetilde{Z}_m^{(2)}/\sqrt{m} \geq 0 > L^{(2)}\right) \\
=&\int_0^\infty \int_0^\infty \int_{-z_1/\sqrt{m}}^0 \int_{-z_2/\sqrt{m}}^0 q_m(y_1, y_2, z_1, z_2) \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}z_1 \mathrm{d}z_2.
\end{aligned}
$$

By Taylor expansion, we have

$$
q_m(y_1, y_2, z_1, z_2) = q_m(0, 0, z_1, z_2) + y_1 \frac{\partial}{\partial y_1} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2) + y_2 \frac{\partial}{\partial y_2} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2), \tag{E.3}
$$

where $\bar{y}_1 \in (-z_1/\sqrt{m}, 0)$ and $\bar{y}_2 \in (-z_2/\sqrt{m}, 0)$. Then

$$
\begin{aligned}
&\mathbb{P}\left(L_m^{(1)} \geq 0 > L^{(1)}, L_m^{(2)} \geq 0 > L^{(2)}\right) \\
=&\int_0^\infty \int_0^\infty \int_{-z_1/\sqrt{m}}^0 \int_{-z_2/\sqrt{m}}^0 \Big[q_m(0, 0, z_1, z_2) + y_1 \frac{\partial}{\partial y_1} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2) \\
&+ y_2 \frac{\partial}{\partial y_2} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2)\Big] \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}z_1 \mathrm{d}z_2 \\
=&\frac{1}{m} \int_0^\infty \int_0^\infty z_1 z_2 q_m(0, 0, z_1, z_2) \mathrm{d}z_1 \mathrm{d}z_2 \\
&+ \int_0^\infty \int_0^\infty \int_{-z_1/\sqrt{m}}^0 \int_{-z_2/\sqrt{m}}^0 \left[y_1 \frac{\partial}{\partial y_1} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2) + y_2 \frac{\partial}{\partial y_2} q_m(\bar{y}_1, \bar{y}_2, z_1, z_2)\right] \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}z_1 \mathrm{d}z_2 \\
\leq&\frac{1}{m} \int_0^\infty \int_0^\infty z_1 z_2 \bar{q}_{0,m}(z_1, z_2) \mathrm{d}z_1 \mathrm{d}z_2 + \int_0^\infty \int_0^\infty \int_{-z_1/\sqrt{m}}^0 \int_{-z_2/\sqrt{m}}^0 (|y_1| + |y_2|) \bar{q}_{1,m}(z_1, z_2) \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}z_1 \mathrm{d}z_2
\end{aligned}
$$

$$=\frac{1}{m}\int_0^\infty\int_0^\infty z_1z_2\bar{q}_{0,m}(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2-\frac{1}{2m^{3/2}}\int_0^\infty\int_0^\infty(z_1^2z_2+z_1z_2^2)\bar{q}_{1,m}(z_1,z_2)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2$$
$$=\mathcal{O}\left(\frac{1}{m}\right).$$

Combine this with (E.1) and (E.2), then we obtain that

$$\mathrm{Var}[\rho_{mn}]=\mathcal{O}\left(\frac{1}{m}\right)+\mathcal{O}\left(\frac{1}{n}\right).$$

E.2 Analysis in D.4.3

$$\begin{aligned}
&\mathbb{E}\left[\bar{D}_1\mathbb{1}\left\{0\le L_m^{(2)}\le 2\pi\epsilon_m\right\}\right]\\
=&\mathbb{E}\left[\left(\mathbb{1}\left\{0\le L_m^{(1)}\le 2\pi\epsilon_m\right\}-\mathbb{1}\left\{-2\pi\epsilon_m< L_m^{(1)}<0\right\}\right)\mathbb{1}\left\{0\le L_m^{(2)}\le 2\pi\epsilon_m\right\}\right]\\
=&\mathbb{E}\left[\mathbb{1}\left\{0\le L_m^{(1)}\le 2\pi\epsilon_m\right\}\mathbb{1}\left\{0\le L_m^{(2)}\le 2\pi\epsilon_m\right\}\right]-\mathbb{E}\left[\mathbb{1}\left\{-2\pi\epsilon_m< L_m^{(1)}<0\right\}\mathbb{1}\left\{0\le L_m^{(2)}\le 2\pi\epsilon_m\right\}\right].
\end{aligned}\tag{E.4}$$

We examine the two terms on the RHS of (E.4). By Taylor expansion (E.3), it follows that

$$\begin{aligned}
&\text{The first term on the RHS of (E.4)}\\
=&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-z_1/\sqrt{m}}^{2\pi\epsilon_m-z_1/\sqrt{m}}\int_{-z_2/\sqrt{m}}^{2\pi\epsilon_m-z_2/\sqrt{m}}q_m(y_1,y_2,z_1,z_2)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2\\
=&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-z_1/\sqrt{m}}^{2\pi\epsilon_m-z_1/\sqrt{m}}\int_{-z_2/\sqrt{m}}^{2\pi\epsilon_m-z_2/\sqrt{m}}\left[q_m(0,0,z_1,z_2)+y_1\frac{\partial}{\partial y_1}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)\right.\\
&\left.+y_2\frac{\partial}{\partial y_2}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)\right]\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2\\
=&(2\pi\epsilon_m)^2\int_{\mathbb{R}}\int_{\mathbb{R}}q_m(0,0,z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2+\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-z_1/\sqrt{m}}^{2\pi\epsilon_m-z_1/\sqrt{m}}\int_{-z_2/\sqrt{m}}^{2\pi\epsilon_m-z_2/\sqrt{m}}\\
&\left[y_1\frac{\partial}{\partial y_1}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)+y_2\frac{\partial}{\partial y_2}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)\right]\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2\\
=&(2\pi\epsilon_m)^2\int_{\mathbb{R}}\int_{\mathbb{R}}q_m(0,0,z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2+\mathcal{O}\left(\epsilon_m^3+\frac{\epsilon_m^2}{\sqrt{m}}\right),
\end{aligned}$$

where the last equality is because by Assumption E.1,

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-z_1/\sqrt{m}}^{2\pi\epsilon_m-z_1/\sqrt{m}}\int_{-z_2/\sqrt{m}}^{2\pi\epsilon_m-z_2/\sqrt{m}}\left[y_1\frac{\partial}{\partial y_1}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)+y_2\frac{\partial}{\partial y_2}q_m(\bar{y}_1,\bar{y}_2,z_1,z_2)\right]\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2\\
\le&\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-z_1/\sqrt{m}}^{2\pi\epsilon_m-z_1/\sqrt{m}}\int_{-z_2/\sqrt{m}}^{2\pi\epsilon_m-z_2/\sqrt{m}}(|y_1|+|y_2|)\bar{q}_{1,m}(z_1,z_2)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2\\
=&\int_{\mathbb{R}}\int_{\mathbb{R}}\left[(2\pi\epsilon_m)^3-\frac{z_1+z_2}{\sqrt{m}}(2\pi\epsilon_m)^2\right]\bar{q}_{1,m}(z_1,z_2)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}z_1\mathrm{d}z_2
\end{aligned}$$

$$=\mathcal{O}\left(\epsilon_m^3+\frac{\epsilon_m^2}{\sqrt{m}}\right).$$

Similarly, it can be shown that

$$\text{The second term on the RHS of (E.4)} = (2\pi\epsilon_m)^2\int_{\mathbb{R}}\int_{\mathbb{R}} q_m(0,0,z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2+\mathcal{O}\left(\epsilon_m^3+\frac{\epsilon_m^2}{\sqrt{m}}\right).$$

Hence, we have

$$\text{(E.4)}=\mathcal{O}\left(\epsilon_m^3+\frac{\epsilon_m^2}{\sqrt{m}}\right).$$

Therefore, by Assumption D.2 and (D.1), it follows that $\left|\sigma_{mn}^{-2}\cdot\mathbb{E}[\bar{D}_1\bar{D}_2]\right|=\mathcal{O}\left(\sigma_{mn}^{-2}\cdot(\epsilon_m^3+\epsilon_m^2/\sqrt{m})\right)\to 0$ as $m\to\infty$.

E.3 Analysis in D.4.4

$$\begin{aligned}
&\mathbb{E}[\widetilde{D}_1\widetilde{D}_2]\\
=&\mathbb{E}\left\{\left[\frac{L_m^{(1)}}{\epsilon_m}-\sin\left(\frac{L_m^{(1)}}{\epsilon_m}\right)\right]\mathbb{1}\left\{|L_m^{(1)}|\le 2\pi\epsilon_m\right\}\cdot\left[\frac{L_m^{(2)}}{\epsilon_m}-\sin\left(\frac{L_m^{(2)}}{\epsilon_m}\right)\right]\mathbb{1}\left\{|L_m^{(2)}|\le 2\pi\epsilon_m\right\}\right\}\\
=&\epsilon_m^2\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi}(u-\sin u)(v-\sin v)q_m(\epsilon_m u-z_1/\sqrt{m},\epsilon_m v-z_2/\sqrt{m},z_1,z_2)\mathrm{d}u\mathrm{d}v\mathrm{d}z_1\mathrm{d}z_2\\
=&\epsilon_m^2\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi}(u-\sin u)(v-\sin v)\left[q_m(0,0,z_1,z_2)\right.\\
&\left.+(\epsilon_m u-z_1/\sqrt{m})\frac{\partial}{\partial y_1}q_m(\bar{u},\bar{v},z_1,z_2)+(\epsilon_m u-z_2/\sqrt{m})\frac{\partial}{\partial y_2}q_m(\bar{u},\bar{v},z_1,z_2)\right]\mathrm{d}u\mathrm{d}v\mathrm{d}z_1\mathrm{d}z_2\\
=&\epsilon_m^2\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi}(u-\sin u)(v-\sin v)\left[(\epsilon_m u-z_1/\sqrt{m})\frac{\partial}{\partial y_1}q_m(\bar{u},\bar{v},z_1,z_2)\right.\\
&\left.+(\epsilon_m u-z_2/\sqrt{m})\frac{\partial}{\partial y_2}q_m(\bar{u},\bar{v},z_1,z_2)\right]\mathrm{d}u\mathrm{d}v\mathrm{d}z_1\mathrm{d}z_2\\
=&\epsilon_m^2\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{-2\pi}^{2\pi}\int_{-2\pi}^{2\pi}(2\pi)^2\left[(4\pi\epsilon_m+(|z_1|+|z_2|)/\sqrt{m})\frac{\partial}{\partial y_1}\bar{q}_{1,m}(z_1,z_2)\right]\mathrm{d}u\mathrm{d}v\mathrm{d}z_1\mathrm{d}z_2\\
=&\mathcal{O}\left(\epsilon_m^3+\frac{\epsilon_m^2}{\sqrt{m}}\right).
\end{aligned}$$

Combine with (D.1), then it follows that $\left|\sigma_{mn}^{-2}\cdot\mathbb{E}[\widetilde{D}_1\widetilde{D}_2]\right|=\mathcal{O}\left(\sigma_{mn}^{-2}\cdot(\epsilon_m^3+\epsilon_m^2/\sqrt{m})\right)\to 0$ as $m\to\infty$.