Advances in Derivative Estimation:

Volk-Makarewicz, W.M.

2014

document version
Publisher's PDF, also known as Version of record

Chapter 3

Supplemental Analysis of MVD Ranked-Data Estimators

3.1 Introduction

From the results of the numerical experiments comparing the *Infinitesimal Perturbation Analysis* (IPA) and *Measure Valued Differentiation* (MVD) ranked-data estimators in Chapter 2, the most unexpected result is that the variance of the IPA estimator is significantly less than that of the MVD counterparts. In general praxis, as mentioned in Section 1.2.3, the variances of IPA and MVD derivative estimators are comparable, with MVD compensating for increased computational time with improved variance. The IPA derivative estimation method is especially effective for ranked-data statistics because the pathwise derivative is between two order statistics as shown in the previous chapter.

As discussed in Section 1.2.2.2 a standard technique for reducing the variance of the MVD estimator is to positively correlate the performance functions $Z^+ = h(X^+)$ and $Z^- = h(X^-)$. For moments and probabilities this yields a significant decrease in the variance of the MVD derivative estimation of these statistics. However, from earlier unpresented simulations for ranked-data estimation, MVD conferred minimal variance reduction from correlating the input random variables. This chapter is an extended analysis, following Chapter 2, of the simulation performance of the MVD ranked-data derivative estimators given our assumption of k samples of m i.i.d. random variables $\mathbf{Z} = (Z_i) : 1 \le i \le m)$. Our focus is on the quantile problem though this analysis equally applies to the order statistic problem. The forthcoming paragraph is a reminder to the reader of the ranked-data statistic and the MVD derivative estimators.

The sequence of random variables $\mathbf{Z}$ are m copies of a random variable $Z = h(X)$ with distribution F_θ, $\theta \in \Theta \subset \mathbb{R}$. The ranked-data statistic for the order statistic problem is denoted by $Z_{l:m}$, $1 \le l \le m$, and pertinently, by the order statistic $Z_{\lceil \alpha m \rceil :m}$, $\alpha \in (0,1)$ for the quantile problem. The parameter θ represents a controllable parameter of the random vector of inputs X. We assume that h is only a mapping of θ via the input vector. The quantile function of F_θ is denoted by $q_\alpha(\theta)$. From Equation (2.4), we will again use the $\text{ord}_{\alpha,m}$ notation to alternatively represent the order statistic associate to the α-level quantile consisting of

m samples. This notation focuses especially on the random variables that form the order statistic. This is needed because for MVD ranked-data estimators the random variables are not identically distributed, and also needed for the various order statistics that occur in the derivations. The $\mathrm{ord}_{\alpha,m}$ representation for the i.i.d. order statistic is repeated below:

$$Z_{\lceil \alpha m \rceil:m} = \mathrm{ord}_{\alpha,m}(Z_1, \ldots, Z_m).$$

In this alternative notation, with the estimator being a sample mean from k sequences, the MVD standard, $l = 1$, and symmetric, $l = 2$, ranked-data derivative estimators have the respective descriptions:

$$D_l^{\mathrm{MVD}}(m,k) = \frac{1}{k}\sum_{j=1}^{k} D_{l,j}(m). \quad l = 1, 2,$$

where

$$\begin{aligned} D_{1,j}(m) = c_\theta \sum_{i=1}^{m} \big(&\mathrm{ord}_{\alpha,m}\big(Z_1(j), \ldots, Z_{i-1}(j), Z_i^+(j), Z_{i+1}(j), \ldots, Z_m(j)\big) \\ &- \mathrm{ord}_{\alpha,m}\big(Z_1(j), \ldots, Z_{i-1}(j), Z_i^-(j), Z_{i+1}(j), \ldots, Z_m(j)\big)\big), \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} D_{2,j}(m) = mc_\theta \big(&\mathrm{ord}_{\alpha,m}\big(Z_1(j), \ldots, Z_{m-1}(j), Z_m^+(j)\big) \\ &- \mathrm{ord}_{\alpha,m}\big(Z_1(j), \ldots, Z_{m-1}(j), Z_m^-(j)\big)\big). \end{aligned} \tag{3.2}$$

The constant $c_\theta > 0$ is the pre-factor for the measure-valued derivative of Z.

The core aspect of this chapter are the results for the asymptotic variance of both derivative estimators in terms of the sample size m. The form of these expressions differ due to the dependence within measure-valued derivative random variable pair (Z^+, Z^-) and its dependence with its associate random variable Z within the sequence. These results form the basis of our analysis. The main corollary we obtain, given mutual independence of all random variables, is that the variance of the standard derivative estimator, asymptotically as $m \to \infty$, is less than that of the symmetric estimator. We also show that the Hahn-Jordan choice of measure-valued derivative has least variance for both ranked-data estimators. We also investigate variance reduction by importance sampling. The consolidation of the results leads to an understanding why the muted variance reduction is discovered.

This chapter has three parts. The story and insight from this chapter is in Section 3.2. In Section 3.2.1 we discuss the dependence structures between Z

and the associate random variables (Z^+, Z^-), and the main assumptions that are needed for all of the proofs. The results are presented in Sections 3.2.2 - 3.2.6. In these sections, we also discuss the key aspects in the derivation of these results and those that assist subsequent results. In Section 3.2.2, we begin by pondering further over the MVD strong consistency result from the previous chapter. This subsection also contains key computations and relations that re-occur in the later three subsections and in the proofs. The asymptotic variance computations are presented where the limit is taken w.r.t. the number of samples m, for the standard and symmetric estimators in Section 3.2.3 and 3.2.4. In Section 3.2.5 we prove for an i.i.d. sequence that asymptotically, as $m \to \infty$, a smaller variance is attained from utilizing the standard estimator and Hahn-Jordan choice of measure-valued derivative distributions. The importance sampling result is asserted in Section 3.2.6. The analysis of our results inferring the simulation performance of MVD derivative estimation of ranked-data is in Section 3.3. The longer derivations for the new results from Section 3.2 are given in Section 3.4, shown respectively in separate subsections.

3.2 Presentation of the Results

3.2.1 Preliminaries

From Section 1.2.2.2, the precision and, by extension, computational speed, of derivative estimation via MVD is determined through the degree of positive correlation between Z^+ and Z^-:

$$\mathrm{Var}_\theta\left(Z^+ - Z^-\right) = \mathrm{Var}_\theta\left(Z^+\right) + \mathrm{Var}_\theta\left(Z^-\right) - \mathrm{Cov}_\theta\left(Z^+, Z^-\right). \tag{3.3}$$

For ranked-data derivative estimation, though, this correlation behaviour does not behave as simply. From the four examples in Sections 2.3.2 and 2.4.2 if the variance behaved according to the above expression then, for instance, there would be a noticeable reduction in variance for the MVD estimators where the random variables Z^+, Z^- are dependent to when they are independently generated: see [53], [54] for examples.

The asymptotic variances are calculated w.r.t. three dependence structures. We note that this delineation to these dependencies between (Z^+, Z^-) and the associate random variable Z form a mutually exclusive and exhaustive division:

- **(a)** Mutual independence between Z, Z^+ and Z^-.
- **(b)** Z^+ and Z^- are mutually dependent, but the pair (Z^+, Z^-) is independent of Z, and

(c) (Z^+, Z^-) is dependent via Z, in which Z^+ and Z^- may or may not be mutually dependent random variables.

Specifically, within a sequence $\mathbf{Z}$, for either dependence structure **(b)** or **(c)**, the dependence, if it is present, is between the random variables Z_i^+ and Z_i^-, $1 \le i \le m$. For $j \ne i$, $1 \le j \le m$, Z_i^+ and Z_j^- are independent of each other. Also, for dependence structure **(c)**, if it is present, the dependence is only between the measure-valued derivative pair (Z_i^+, Z_i^-) and the random variable Z_i. For $j \ne i$, the random variable pair (Z_i^+, Z_i^-) is independent of Z_j.

The results for the asymptotic variance computations are written in terms of the marginal and/or bivariate distributions consisting of the measure-valued derivative pair (Z^+, Z^-), and the original random variable Z. The bivariate distribution conveys the effect that the correlation between the random variables has on the expressions.

For each of the results presented we will use the assumptions from Sections 2.3.1 and 2.4.1. For ease of reference we provide the conditions in the following:

(B1) $\mathbf{Z}(j) = (Z_i(j) : 1 \le i \le m)$, for $1 \le j \le k$, is a family of mutually independent random variables with a continuous distribution function.

(B2) The random variable Z is $\mathfrak{D}$-differentiable with $\mathfrak{D} = \mathfrak{B}_v, \mathfrak{C}_v$, and a version of its weak derivative is given by (c_θ, Z^+, Z^-). For estimators D_l^{MVD}, for $l = 1, 2$, we assume that $(Z_i^+(j), Z_i^-(j))$, for $1 \le i \le m$ and $1 \le j \le k$, are an i.i.d. collection of random variable pairs that are independent of $\mathbf{Z}(l)$, $1 \le l \le k$, except, possibly, $Z_i(j)$.

(B3) The dependence structure within each element of the collection $((Z_i(j), Z_i^+(j), Z_i^-(j)) : 1 \le i \le m, 1 \le j \le k)$, consisting of i.i.d. copies of the random variable triple (Z, Z^+, Z^-) has one one the following three forms:

- **(a)** Mutual independence between Z, Z^+ and Z^-,
- **(b)** Z^+ and Z^- are mutually dependent, but the pair (Z^+, Z^-) is independent of Z, and
- **(c)** (Z^+, Z^-) is dependent via Z, in which Z^+ and Z^- may or may not be mutually dependent random variables.

(B4) The random variable $(Z_i(j))$ is square integrable for $1 \le i \le m$, $1 \le j \le k$.

(B5) In a neighbourhood of $q_\alpha(\theta)$ it holds that $f_\theta(x) > 0$.

(B6) Let Θ_0 be an open neighbourhood of θ_0, where θ_0 denotes the value of θ for which the derivative is evaluated, and let $B(\alpha)$ be an open neighbourhood of $x = q_\alpha(\theta_0)$. Assume that f_θ is continuously differentiable on Θ_0 for all $x \in B(\alpha)$.

The importance sampling lemma requires two additional assumptions and these are supplied in Section 3.2.6. On a final note, when we refer to an equality between two random variables, this is in distribution, unless stated otherwise.

3.2.2 Asymptotic Expectation

This section focuses on key concepts and expressions presented in the MVD component of the Strong Consistency Theorem (Theorem 2.2), and examine this result with regard to the dependence structure of the random variable pair (Z^+, Z^-). The aspects presented here are utilized in the remainder of this section. In addition, these aspects are needed pedagogically as the line of argument for each of the forthcoming derivations follows from this theorem.

We begin by reviewing notation and the key aspects of the MVD strong consistency proof as given in Theorem 2.2. This derivation is valid when the measure-valued derivative random variable pair (Z^+, Z^-) does not depend on the associate random variable Z, i.e., if Assumptions **(B3(a))** and **(B3(b))** hold. However, if Assumption **(B3(c))** is satisfied, the tower property of conditional expectation is also needed, in which each order statistic $Z^+_{\lceil \alpha m \rceil : m}$, $Z^-_{\lceil \alpha m \rceil : m}$ is conditioned w.r.t. the sequence $\mathbf{Z}$ to deal with the possible dependence between Z_m and the random variable Z^+_m or Z^-_m before conditioning w.r.t. $\mathbf{Z} \backslash Z_m = (Z_i, 1 \le i \le m-1)$. While this additional detail is minor, this detail leads bivariate distributions that are not Lebesgue absolute continuous[1].

Let $Z^+_{\lceil \alpha m \rceil : m}$ and $Z^-_{\lceil \alpha m \rceil : m}$, $\alpha \in (0,1)$, be the order statistics consisting of m independent random variables where $m-1$ of these, $\mathbf{Z} \backslash Z_m$ are copies of Z, and the random variable Z_m is respectively substituted by Z^+_m, Z^-_m. Alternatively, these order statistics are written as

$$
\begin{aligned}
Z^+_{\lceil \alpha m \rceil : m} &= \operatorname{ord}_{\alpha,m}\big(Z_1, \ldots, Z_{m-1}, Z^+_m\big), \quad \text{and,} \\
Z^-_{\lceil \alpha m \rceil : m} &= \operatorname{ord}_{\alpha,m}\big(Z_1, \ldots, Z_{m-1}, Z^-_m\big).
\end{aligned} \tag{3.4}
$$

We begin by recapitulating the Strong Consistency Theorem.

[1] Think of a bivariate distribution that also contains atoms or more likely, is a singular distribution. An example is the random vector (X, Y), X and Y are univariate normal distributions and X and Y are realted via an affine transformation.

Theorem 3.1 (Strong Consistency)**.** *Suppose that Assumptions* **(B1)** *to* **(B6)** *are satisfied. Then, for $l = 1,2$, it holds almost surely that*

$$\lim_{(m,k)\to(\infty,\infty)} D_l^{\mathrm{MVD}}(m,k) = \frac{\partial}{\partial\theta} q_\alpha(\theta).$$

For the sequence of random variables $\mathbf{Z}\backslash Z_m$ we recall in the derivation hat $Y_1 = Z_{\lceil\alpha m\rceil-1:m-1}$ and $Y_2 = Z_{\lceil\alpha m\rceil:m-1}$ are order statistics composed from these $m-1$ elements; i.e.,

$$\begin{aligned} Y_1 &= \mathrm{ord}_{\alpha_1,m}(Z_1,\dots,Z_{m-1}), \quad \text{and,} \\ Y_2 &= \mathrm{ord}_{\alpha_2,m}(Z_1,\dots,Z_{m-1}), \end{aligned}$$

where α_1 such that $\lceil\alpha m\rceil - 1 = \lceil\alpha_1(m-1)\rceil$ and α_2 such that $\lceil\alpha m\rceil = \lceil\alpha_2(m-1)\rceil$.

Given either Assumption **(B3(a))** or **(B3(b))** the first aspect that we draw to is the distributional relation between $Z_{\lceil\alpha m\rceil:m}$ and $\mathbf{Z}\backslash Z_m$, when we include Z_m^+ into the sequence $\mathbf{Z}\backslash Z_m$. The outcomes, from the argument preceding Equation (2.9), are either Y_1, Y_2, or Z_m^+. If $Z_m^+ \le Y_1$, $Z^+_{\lceil\alpha m\rceil:m} = Y_1$ as there are $\lceil\alpha m\rceil$ random variables less than or equal to Y_1. If $Z_m^+ \ge Y_2$, $Z^+_{\lceil\alpha m\rceil:m} = Y_2$, since Z_m^+ has a rank greater than $\lceil\alpha m\rceil$, and $Z^+_{\lceil\alpha m\rceil:m} = Z_m^+$ if $Y_1 < Z_m^+ \le Y_2$ as Z_m^+ is the random variable with the $\lceil\alpha m\rceil^{\text{th}}$ smallest realization. Or in terms of probabilities, with F_θ^+ being the distribution function of Z^+, this relation yields

$$\mathbb{E}_\theta\Big[Z^+_{\lceil\alpha m\rceil:m}\Big|\mathbf{Z}\backslash Z_m\Big] = \begin{cases} \textbf{Random Variable} & \textbf{Probability} \\ Y_1 & F_\theta^+(Y_1) \\ \mathbb{E}_\theta\big[Z_m^+ \,\big|\, Z_m^+ \in (Y_1,Y_2]; \mathbf{Z}\backslash Z_m\big] & F_\theta^+(Y_2) - F_\theta^+(Y_1) \\ Y_2 & 1 - F_\theta^+(Y_2). \end{cases} \tag{3.5}$$

A similar expression occurs for $\mathbb{E}[Z^-_{\lceil\alpha m\rceil:m}|\mathbf{Z}\backslash Z_m]$, with $Z^- \sim F_\theta^-$. This identity (or a related identity) is central to the derivations of the asymptotic variance and the importance sampling lemmas, as will be seen in Section 3.4.

From Theorem 2.1, both the symmetric and standard derivative estimators have the same mean, and, after a little computation, this proof reduces to the expression

$$c_\theta \mathbb{E}_\theta\Big[Z^+_{\lceil\alpha m\rceil:m} - Z^-_{\lceil\alpha m\rceil:m}\Big|\mathbf{Z}\backslash Z_m\Big] = -c_\theta \int_{Y_1}^{Y_2} F_\theta^+(z) - F_\theta^-(z)\, dz, \tag{3.6}$$

and this random variable is uniformly integrable, for a proof see Equation (3.16) below. Similar reductions also occur for the other results. More importantly, Equation (3.6) has the same value irrespective of the measure-valued derivative triple $(c_\theta, f_\theta^+, f_\theta^-)$, where f_θ^+, f_θ^- are the density functions of Z^+, Z^-. Indeed, from the definition of the measure-valued derivative, Section 1.2.2.2, Equation (1.21) yields in particular:

$$\begin{aligned} c_\theta \int_{Y_1}^{Y_2} F_\theta^+(z) - F_\theta^-(z)\,dz &= \int_{Y_1}^{Y_2} c_\theta \left(\int_{\mathbb{R}} 1\{x \le z\}\, f_\theta^+(x)\,dx - \int_{\mathbb{R}} 1\{x \le z\}\, f_\theta^-(x)\,dx \right) dz \\ &= \int_{Y_1}^{Y_2} \int_{\mathbb{R}} 1\{x \le z\}\, f_\theta'(x)\,dx\,dz. \end{aligned}$$

To emphasize this statement, if $(\tilde{c}_\theta, \tilde{f}_\theta^+, \tilde{f}_\theta^-)$ denotes the Hahn-Jordan measure-valued derivative triple, then

$$c_\theta \mathbb{E}_\theta\Big[Z^+_{\lceil \alpha m\rceil:m} - Z^-_{\lceil \alpha m\rceil:m} \Big| \mathbf{Z}\backslash Z_m \Big] = \tilde{c}_\theta \mathbb{E}_\theta\Big[\tilde{Z}^+_{\lceil \alpha m\rceil:m} - \tilde{Z}^-_{\lceil \alpha m\rceil:m} \Big| \mathbf{Z}\backslash Z_m \Big], \tag{3.7}$$

in which $\tilde{Z}^+_{\lceil \alpha m\rceil:m}$, $\tilde{Z}^-_{\lceil \alpha m\rceil:m}$ are the analogous versions of the order statistic in Equation (3.4), i.e.,

$$\begin{aligned} \tilde{Z}^+_{\lceil \alpha m\rceil:m} &= \mathrm{ord}_{\alpha,m}\left(Z_1, \dots, Z_{m-1}, \tilde{Z}^+_m\right), \quad \text{and,} \\ \tilde{Z}^-_{\lceil \alpha m\rceil:m} &= \mathrm{ord}_{\alpha,m}\left(Z_1, \dots, Z_{m-1}, \tilde{Z}^-_m\right), \end{aligned}$$

and Z_m is instead respectively substituted by $\tilde{Z}^\pm_m$. These Hahn-Jordan random variables are distributed as $\tilde{Z}^+ \sim \tilde{F}_\theta^+$, $Z^- \sim \tilde{F}_\theta^-$.

The derivation for this proof is identical for both dependence structures **(B3(a))** or **(B3(b))** due to the additivity property of conditional expectations applied to (3.6). For the dependence structure **(B3(c))**, the distribution equivalence in (3.5) is different in which the distribution function are conditioned w.r.t. Z_m. This is due to the possible dependence of the random variables $Z^\pm$ on Z affecting the ranks attained by $Z^\pm$ in the respective order statistics $Z^+_{\lceil \alpha m\rceil:m}$, $Z^-_{\lceil \alpha m\rceil:m}$. The distribution of Z^+_m given Z_m is denoted by $F_\theta^+(\cdot|Z_m)$, and we obtain:

$$\mathbb{E}_\theta\Big[Z^+_{\lceil \alpha m\rceil:m} \Big| \mathbf{Z} \Big] = \begin{cases} \textbf{Random Variable} & \textbf{Probability} \\ Y_1 & F_\theta^+(Y_1|Z_m) \\ \mathbb{E}_\theta\big[Z^+_m \,\big|\, Z^+_m \in (Y_1, Y_2]; \mathbf{Z} \big] & F_\theta^+(Y_2|Z_m) - F_\theta^+(Y_1|Z_m) \\ Y_2 & 1 - F_\theta^+(Y_2|Z_m). \end{cases}$$

A similar result is attained for the order statistic $\mathbb{E}_\theta[Z^-_{\lceil \alpha m\rceil:m}|\mathbf{Z}]$. The equivalent expression to Equation (3.6) contingent on Assumption **(B3(c))** is then

$$c_\theta \mathbb{E}_\theta\Big[Z^+_{\lceil \alpha m\rceil:m} - Z^-_{\lceil \alpha m\rceil:m} \Big| \mathbf{Z} \Big] = -c_\theta \int_{Y_1}^{Y_2} F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\,dz. \tag{3.8}$$

The above expression, via Lemma A.2, is uniformly integrable by the reasoning observed in Equation (3.16). A combination of the tower property of conditional expectation and Fubini's Theorem then yields Equation (3.6):

$$\begin{aligned}\mathbb{E}_\theta\left[\int_{Y_1}^{Y_2} F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\,dz\,\middle|\,\mathbf{Z}\backslash Z_m\right] \\ = \int_{Y_1}^{Y_2} \mathbb{E}_\theta\big[F_\theta^+(z|Z_m)\,\big|\,\mathbf{Z}\backslash Z_m\big] - \mathbb{E}_\theta\big[F_\theta^-(z|Z_m)\,\big|\,\mathbf{Z}\backslash Z_m\big]\,dz \\ = \int_{Y_1}^{Y_2} F_\theta^+(z) - F_\theta^-(z)\,dz.\end{aligned}$$

The motivation to consider functions that are not absolutely continuous w.r.t. the Lebesgue measure is that one of the measure-valued derivatives will have the same distribution as the associate random variable; for example, the derivative w.r.t. the rate parameter for the exponential distribution, Section 1.2.4.1 where $Z^+ = Z$. To attain the distributions F_θ^+, F_θ^- contingent on Assumption **(B3(c))** we need to consider the related conditional distributions in the form of an Lebesgue-Stieltjes integral, [96], p .158. This is permitted as distribution functions are evidently finite for all bounded intervals. As we are conducting our computations via the Lebesgue integral, we hereafter simply refer in this chapter to the Lebesgue-Stieltjes integral as the Stieltjes integral and similarly for other associate concepts.

In the treatment of Equation (3.8), we need to be mindful, technically, of the possibility in the present formulation that the conditional distributions $F_\theta^+(\cdot|Z_m)$, $F_\theta^-(\cdot|Z_m)$ are actually random variables. Following the argument on p. 226 in [96], however, these random variables, $F_\theta^+(\cdot|Z_m)$, $F_\theta^-(\cdot|Z_m)$, may not be probability measures as there may be a non-zero probability that the additive probability property is violated. For these particular conditional distributions, the possibility may arise when integration w.r.t. Z_m is being conducted that the region of integration is a countable union of disjoint sets. Though the elements ω for that the countable additivity property does not hold for a particular countable union of such sets has measure zero, there is an uncountable union of such possibilities and the total measure may not be zero.

With this in mind, from this point on in this chapter we assume that $F_\theta^+(\cdot|Z_m)$, $F_\theta^+(\cdot|Z_m)$ are *regular conditional distributions*, [62], for the random variables Z_m^+, Z_m^-. Regular conditional distributions, and more generally *regular conditional probabilities*, are almost sure variants of the earlier conditional distributions but with the requirement that $F_\theta^+(\cdot|Z_m)$, $F_\theta^+(\cdot|Z_m)$, are probability measures for any realization Z_m. From [62] p. 107, existence and almost everywhere uniqueness of the regular conditional distribution is assured for any random variable or random element.

Apart from the form of the random variables and the measurability of conditional distribution functions, we also need to consider how the conditional distributions arose. From the MVD component of the derivation in Theorem 2.2, the distribution functions in Equation (3.8) are attained via the integration by parts formula. However, since the conditional random variables are not necessarily Lebesgue absolute continuous, we only can be certain that the regular conditional distributions possess the cádlág[2] property. The difference in expression to (3.8) due to the corresponding Stieltjes integration by parts formula, [96] p. 206, taking into account this property is that the argument for both regular conditional distributions is now the left limit[3], i.e. $F_\theta^+(z^-|Z_m)$. In integral form, the left limit is equivalent to an integral defined on an open set, i.e.

$$F_\theta^+(z^-|Z_m) = \int_{(-\infty,z)} dF_\theta^+(u|Z_m),$$

with a similar expression for $F_\theta^-(z^-|Z_m)$. This is just a technical nuance as, ultimately, the unconditional distributions F_θ^+, F_θ^- are continuous in the neighbourhood of the quantile by Assumption **(B6)**. This result is attained via a conditional expectation, conditioning w.r.t. the sequence $\mathbf{Z}\backslash Z_m$, and we derive this result below for the distribution function F_θ^+. The conditional expectation w.r.t. the sequence $\mathbf{Z}\backslash Z_m$ results in a bivariate Stieltjes integral w.r.t. G_θ^+ denoting the bivariate distribution of the random variable pair (Z^+, Z). From the Disintegration Theorem, [62] p. 108, as $|G_\theta^+| \leq 1$, we can then decompose the accompanying Stieltjes measure in the intuitive manner $dG_\theta^+(u,v) = dF_\theta(v|u)dF_\theta^+(u)$ and

$$\begin{aligned}
\mathbb{E}_\theta\big[F_\theta^+(z|Z_m)\big|\mathbf{Z}\backslash Z_m\big] &= \int_{(-\infty,z)\times\mathbb{R}} dG_\theta^+(u,v) \\
&= \int_{(-\infty,z)} \int_{\mathbb{R}} dF_\theta(v|u)\, dF_\theta^+(u) \\
&= \int_{(-\infty,z)} dF_\theta^+(u) \;=\; F_\theta^+(z^-). \qquad (3.9)
\end{aligned}$$

3.2.3 Asymptotic Variance: Standard Estimator

The main concept here is that the asymptotic variance for the standard estimator is not affected by the correlation present within the random vector (Z^+, Z^-).

[2] namely a distribution function that is right continuous and always has a left limit.

[3] the other function in the integration by parts formula $G(z) = z$ is continuous and hence there is no additional sum for when both terms jump at a common value

Actually, the variance reduction only occurs within the sub-leading term in terms of the variance component. This further discussed in Section 3.2.4, where this term is the main component. The exact result is contingent whether or not the random variables (Z^+, Z^-) are dependent via Z, and this is also surprising from Equation (3.3). This dependence either reduces or increases the variance of the leading term and is due to the rank of the removed random variable Z_i, $1 \le i \le m$, influencing the rank of the inserted measure-valued derivative version.

When there is no dependence exhibited, as we will show in part (i) of the lemma, we observe that the asymptotic variance as only a function of the quantile sensitivity. Let G_θ^+ denote the bivariate distribution function for (Z^+, Z) and similarly G_θ^- w.r.t. the random variable pair (Z^-, Z). The result in part (ii) is the dependent analogue to part (i). In fact, setting $G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) = F_\theta^+(q_\alpha(\theta)) \cdot F_\theta(q_\alpha(\theta)) = \alpha F_\theta^+(q_\alpha(\theta))$ and equivalently for $G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))$, the expression for part (i) is attained immediately.

Lemma 3.1 (Asymptotic Variance, Standard Estimator)**.** *Suppose that Assumptions* **(B1)** - **(B2)** *and* **(B4)** - **(B6)** *holds. In addition,*

(i) *if either Assumption* **(B3(a))** *or* **(B3(b))** *is satisfied, as $m \to \infty$ for $k \ge 1$*

$$\lim_{m\to\infty} \mathrm{Var}_\theta\big(D_1^{\mathrm{MVD}}(m,k)\big) = \frac{1}{k}(\alpha^2 + (1-\alpha)^2)\left(\frac{\partial}{\partial\theta} q_\alpha(\theta)\right)^2,$$

(ii) *if Assumption* **(B3(c))** *is satisfied, as $m \to \infty$ for $k \ge 1$*

$$\begin{aligned}
&\lim_{m\to\infty} \mathrm{Var}_\theta\big(D_1^{\mathrm{MVD}}(m,k)\big) \\
&= \frac{c_\theta^2}{k f_\theta^2(q_\alpha(\theta))}\Big(\big(G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) - G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))\big)^2 \\
&\quad + \Big(\big(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\big) - \big(G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) - G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))\big)\Big)^2\Big).
\end{aligned}$$

The method of proof is identical to both parts and so in Section 3.4.1 we derive part (ii), noting the equivalent for part (i) at certain points. As shorthand, we let $B_1^+(m)$, $B_1^-(m)$ denote the order statistics $Z^+_{\lceil\alpha m\rceil:m}$, $Z^-_{\lceil\alpha m\rceil:m}$ in (3.4). This is the form of the notation that is used in the asymptotic variance derivations. The "1" denotes the substitution of the random variable Z_m with $Z_m^\pm$. Similarly, $B_2^\pm(m)$ is defined to be the same order statistic but the random variable Z_{m-1} is substituted for $Z_{m-1}^\pm$:

$$B_2^\pm(m) = \mathrm{ord}_{\alpha,m}\big(Z_1, \dots, Z_{m-2}, Z_{m-1}^\pm, Z_m\big). \tag{3.10}$$

The leading term in this computation is the covariance, $\mathrm{Cov}_\theta(B_1^+(m) - B_1^-(m), B_2^+(m) - B_2^-(m))$. After some computation and use of the tower property of conditional expectation, we attain this representation:

$$\begin{aligned}
&\mathrm{Cov}_\theta\big(B_1^+(m) - B_1^-(m), B_2^+(m) - B_2^-(m)\big)\\
&\quad = \mathbb{E}\big[\mathbb{E}\big[\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\big]\,\mathbb{E}_\theta\big[B_2^+(m) - B_2^-(m)\,\big|\,\mathbf{Z}\big]\,\big|\,\mathbf{Z}\backslash(Z_{m-1}, Z_m)\big]\big]\\
&\quad\quad - \mathbb{E}^2\big[\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big]\big]\,. \qquad (3.11)
\end{aligned}$$

The proof of this computation is given in Section 3.4.1. As discussed in Section 3.2.2, we need to consider distribution functions that are not absolutely continuous w.r.t. the Lebesgue measure, specifically when the random variables Z^+, Z^-, individually depend on Z. We employ the Stieltjes integral in the second half of this derivation to attain the result of part (ii) in the above Lemma.

The reason that the asymptotic variance for this estimator is decoupled from the correlation between Z^+ and Z^- is that in the first term the random variables depicted as conditional expectations in (3.11) given the σ-field induced by the random sequence $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$ are independent of each other. For instance, in the conditional expectation $\mathbb{E}_\theta[B_1^+(m)B_2^-(m)|\mathbf{Z}\backslash(Z_{m-1}, Z_m)]$ the only unconditioned random variables are Z_{m-1}^-, Z_{m-1}, Z_m^+, and Z_m, which are independent by Assumption **(B2)** being from different measure-valued derivative pairs.

3.2.4 Asymptotic Variance: Symmetric Estimator

Inducing correlation between each pair (Z^+, Z^-) does affect the value of the asymptotic variance for this derivative estimator. The extent of this effect depends on the covariance of the measure-valued derivative distribution functions F_θ^+, F_θ^-, with positive correlation resulting in a decrease in variance. In addition, the asymptotic variance does not depend on the associate random variable Z.

For these computations we need two added pieces of notation. For the measure-valued derivative pair (Z^+, Z^-) we denote the distribution function by $F_\theta^\pm$ and the density function by $f_\theta^\pm$. Additionally, for the triple (Z^+, Z^-, Z) the distribution function is depicted as $G_\theta^\pm$ and the density function via $g_\theta^\pm$.

For part (ii) of this lemma, we again require the Stieltjes integral when considering the possible dependencies for both the random variable pair (Z^+, Z^-) and triple (Z^+, Z^-, Z). For the pair (Z^+, Z^-), it is possible that this pair can be generated by a single common random variable and so the density function possess a one-dimensional support on $\mathbb{R}^2$. An example of this is the parameter derivative w.r.t. the mean for the normal distribution, Section 1.2.4.2, where the Rayleigh distributed measure-valued derivatives on either side of a half line are generated by a single $(0,1)$-uniform random variable. The random variable triple (Z^+, Z^-, Z) can be defined via a two-dimensional subspace due to the possibility of Z^+ or Z^- equalling Z, as noted in Section 3.2.2, or that Z^+ and Z^- can be jointly generated by the same common random variables. Hypothetically, the triple (Z^+, Z^-, Z) can also be defined only in a one-dimensional subspace of $\mathbb{R}^3$.

Lemma 3.2 (Asymptotic Variance, Symmetric Estimator)**.** *Suppose that Assumptions* **(B1)** - **(B2)** *and* **(B4)** - **(B6)** *hold. In addition,*

(i) *if Assumption* **(B3(a))** *is satisfied, as $m \to \infty$ for all $k \geq 1$*

$$\begin{aligned}
&\lim_{m\to\infty} \mathrm{Var}_\theta\big(D_2^{\mathrm{MVD}}(m,k)\big) \\
&\quad = \frac{c_\theta^2}{k f_\theta(q_\alpha(\theta))}\Big(2\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right) - \left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right)^2\Big) \\
&\quad := \frac{c_\theta^2}{k f_\theta(q_\alpha(\theta))}\mathrm{Var}_\theta(D_2^{\mathrm{IND}}),
\end{aligned}$$

(ii) *if Assumption* **(B3(b))** *or* **(B3(c))** *is satisfied, as $m \to \infty$ for all $k \geq 1$*

$$\begin{aligned}
&\mathrm{Var}_\theta\left(D_2^{\mathrm{MVD}}(m,k)\right) \\
&\quad = \frac{c_\theta^2}{k f_\theta(q_\alpha(\theta))}\Big(\mathrm{Var}_\theta(D_2^{\mathrm{IND}}) - 4\left(F_\theta^{\pm}(q_\alpha(\theta), q_\alpha(\theta)) - F_\theta^+(q_\alpha(\theta))F_\theta^-(q_\alpha(\theta))\right)\Big).
\end{aligned}$$

For each of the three dependence structures, the derivation begins by noting that the estimator, Equation (3.2), is formed from k i.i.d. sequences. With $B^+(m)$ being a compact description of $Z^+_{\lceil \alpha m\rceil:m}$, and $B^-(m)$ for $Z^-_{\lceil \alpha m\rceil:m}$, see Equation (3.4):

$$\begin{aligned}
\mathrm{Var}_\theta\left(D_2^{\mathrm{MVD}}(m,k)\right) &= \frac{c_\theta^2}{k}\mathrm{Var}_\theta\left(D_{2,1}(m)\right) \\
&= \frac{c_\theta^2}{k}\Big(\mathbb{E}\Big[\mathbb{E}_\theta\Big[m^2\left(B^+(m) - B^-(m)\right)^2\Big|\mathbf{Z}\backslash Z_m\Big]\Big] \\
&\quad - \mathbb{E}^2\big[\mathbb{E}_\theta\big[m\left(B^+(m) - B^-(m)\right)\big|\mathbf{Z}\backslash Z_m\big]\big]\Big). \qquad (3.12)
\end{aligned}$$

The substituted random variable in the sequence $\mathbf{Z}$ for a measure-valued derivative random variable is Z_m. The difference between the three derivations is the conditional expectation $\mathbb{E}[B^+(m)B^-(m)|\mathbf{Z}\backslash(Z_{m-1}, Z_m)]$. If Z^+ and Z^- are independent for each pair, than this term separates by independence. However, if Z^+ and Z^- are dependent, there are nine equivalent possible outcomes; given by a product of three outcomes per order statistic. These outcomes adhere to the same argument that precede Equation (3.5), and a graphical representation of these outcomes is given in Figure 3.1. In particular, the realizations labelled in each of the regions in this figure denote the almost sure outcomes of the random variable $Z^+_{\lceil \alpha m\rceil:m} Z^-_{\lceil \alpha m\rceil:m} | \mathbf{Z} \backslash Z_m$ depending on the realization of the random variable pair (Z^+_m, Z^-_m) in $\mathbb{R}^2$ represented by the axes. Each outcome can be read by cross-referencing to the hypothetical values of the measure-valued derivative pair. The proofs to attain the asymptotic variances for each of these dependence structures is given in Section 3.4.2.

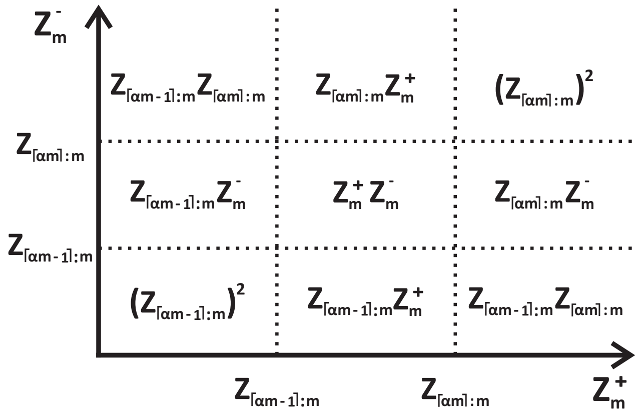

Figure 3.1: The nine equivalent almost sure outcomes for the random variable $Z^+_{\lceil \alpha m\rceil:m} Z^-_{\lceil \alpha m\rceil:m} | \mathbf{Z} \backslash Z_m$.

3.2.5 Comparisons of Variance

The previous two sections primarily answered the effect of dependence between the pair (Z^+, Z^-). This section answers two other questions, focusing on independent behaviour between these two random variables. Firstly, we determine that the theoretical analysis does affirm the simulation evidence in Chapter 2 that the standard derivative estimator yields less variance than the symmetric estimator. Secondly, we analyze the measure-valued derivative choices for both derivative estimators, confirming that the Hahn-Jordan choice of derivative has least variance, and especially the extent of this reduction.

The verification that the standard estimator has reduced variance is based on the asymptotic result. For these results, the standard derivative estimator has subscript $l = 1$, and the symmetric estimator $l = 2$.

Corollary 3.1. *Suppose that Assumption* **(B1)**, **(B2)**, **(B4)** *to* **(B6)** *are satisfied. In* *addition, if dependence structure* **(B3(a))** *holds, then as $m \to \infty$ for all k*

$$\lim_{m\to\infty} \mathrm{Var}_\theta\big(D_1^{\mathrm{MVD}}(m,k)\big) < \lim_{m\to\infty} \mathrm{Var}_\theta\left(D_2^{\mathrm{MVD}}(m,k)\right).$$

Proof:

$$\begin{aligned}
&\lim_{m\to\infty} k f_\theta^2(q_\alpha(\theta))\big(\mathrm{Var}_\theta(D_2^{\mathrm{MVD}}(m,k)) - \mathrm{Var}_\theta(D_1^{\mathrm{MVD}}(m,k))\big)\\
&\quad = c_\theta^2\Big(2\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right) - \left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right)^2\\
&\qquad - (\alpha^2 + (1-\alpha)^2)\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right)^2\Big)\\
&\quad > 2c_\theta^2\left(F_\theta^+(q_\alpha(\theta))\left(1 - F_\theta^+(q_\alpha(\theta))\right) + F_\theta^-(q_\alpha(\theta))\left(1 - F_\theta^-(q_\alpha(\theta))\right)\right) \geq 0,
\end{aligned}$$

since $\alpha^2 + (1-\alpha)^2 < 1$ for $\alpha \in (0,1)$. □

We can investigate this difference by taking the ratio of these two asymptotic variances. The extent of this difference can be estimated via the empirical distribution function of both F_θ^+, F_θ^-. This result requires the complementary distribution function for both Z^+ and Z^-, written as $\bar{F}_\theta^+$, $\bar{F}_\theta^-$ respectively.

$$\begin{aligned}
&\frac{\mathrm{Var}_\theta(D_2^{\mathrm{MVD}}(m,k))}{\mathrm{Var}_\theta(D_1^{\mathrm{MVD}}(m,k))}\\
&\quad = \frac{\left(\alpha^2 + (1-\alpha)^2\right)\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right)^2}{2\Big(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\Big) - \Big(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\Big)^2}\\
&\quad = \left(\alpha^2 + (1-\alpha)^2\right)\left(\frac{F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))}{F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))}\right)\left(\frac{\bar{F}_\theta^-(q_\alpha(\theta)) - \bar{F}_\theta^+(q_\alpha(\theta))}{\bar{F}_\theta^-(q_\alpha(\theta)) + \bar{F}_\theta^+(q_\alpha(\theta))}\right).
\end{aligned}$$

This expression has a maximum at $\alpha^2 + (1-\alpha^2)$ when $F_\theta^+(q_\alpha(\theta)) = 0, 1$ and $F_\theta^-(q_\alpha(\theta)) = 1 - F_\theta^+(q_\alpha(\theta))$. However, in practical cases this exact occurrence is uncommon.

The only term that attains a minimum value from either of the variance expressions, Equations (3.11) and (3.12), via the Hahn-Jordan choice of measure-valued derivative is the second moment $\mathbb{E}_\theta[(Z^+_{\lceil\alpha m\rceil:m} - Z^-_{\lceil\alpha m\rceil:m})^2 | \mathbf{Z}\backslash Z_m]$. In the following equation, involving measures, we assume that Z_m is the random variable removed, replaced respectively by Z_m^+ and Z_m^-. Then denoting $b_m(z)$ for the map of the order statistic conditioned on the random variables $\mathbf{Z}\backslash Z_m$:

$$\begin{aligned}
c_\theta^2\,\mathbb{E}_\theta\left[\left(Z^+_{\lceil\alpha m\rceil:m} - Z^-_{\lceil\alpha m\rceil:m}\right)^2\middle|\mathbf{Z}\backslash Z_m\right] &= c_\theta^2\int_{\mathbb{R}^2}(b_m(z) - b_m(w))^2\,\mathbb{P}_\theta^+(dz)\,\mathbb{P}_\theta^-(dw)\\
&= \int_{\mathbb{R}^2}(b_m(z) - b_m(w))^2\,\mu_\theta^+(dz)\,\mu_\theta^-(dw)\\
&= \int_{\mathbb{R}^2}(b_m(z) - b_m(w))^2\,\left(\tilde{\mu}_\theta^+ + \nu\right)(dz)\left(\mu_\theta^- + \nu\right)(dw)
\end{aligned}$$

$$> \tilde{c}_\theta^2 \int_{\mathbb{R}^2} (b_m(z) - b_m(w))^2 \tilde{\mathbb{P}}_\theta^+(dz) \tilde{\mathbb{P}}_\theta^-(dw).$$

From Section 1.2.2.2, the finite signed measure denoting the measure-valued derivative μ'_θ has a Hahn-Jordan decomposition $\mu'_\theta = \tilde{\mu}_\theta^+ - \tilde{\mu}_\theta^- = \tilde{c}_\theta(\mathbb{P}_\theta^+ - \mathbb{P}_\theta^-)$. And any measure-value derivative $(c_\theta, \mathbb{P}_\theta^+, \mathbb{P}_\theta^-)$ can be attained from the Hahn-Jordan decomposition via $\mu_\theta^\pm = \tilde{\mu}_\theta^\pm + \nu$, before normalizing, where ν is a positive measure.

Especially, from Equation (3.7), the conditional expectation is equal for all choices of measure-valued derivative triple. For a finite sample size m, from Equation (3.12), the Hahn-Jordan choice of measure-valued derivative provides least variance for the symmetric estimator as the leading term for the variance is the second moment. However, for the standard derivative estimator, the second moment is sub-leading of order $O(m^{-1})$ to the covariance term. Via conditioning, the covariance term in (3.11) can be solely expressed by the expectations in (3.7) and so is equal for any choice of f_θ^+, f_θ^-, reverting to density functions. Consequently, even though this choice provides least variance, the effect of the Hahn-Jordan choice of density function $\tilde{f}_\theta^+$, $\tilde{f}_\theta^-$ for reducing the variance or the standard estimator is small. This is the reason why the asymptotic variance of the standard estimator is independent of the choice of the distribution for Z^+ and Z^-. This explanation leads to the corollary:

Corollary 3.2. *If Assumptions* **(B1)**, **(B2)**, **(B4)** *to* **(B6)** *are satisfied, and given mutual independence, Assumption* **(B3(a))**, *for any choice of finite m and k, the Hahn-Jordan choice of measure-valued derivative provides the least variance for both standard and symmetric derivative estimators.*

To end this section, we verify that the choice of Hahn-Jordan distribution functions $\tilde{F}_\theta^+$, $\tilde{F}_\theta^+$ yields least variance for the symmetric estimator in the asymptotic case. For this observation, we let $(c_\nu, F_\nu^+, F_\nu^-)$ be the measure-valued derivative triple associated with the positive measure ν. Defining
$\mathrm{Var}\big(\tilde{D}_2^{\mathrm{MVD}}(k,m)\big)$ as the Hahn-Jordan asymptotic variance, for any other measure-valued derivative triple $(c_\theta, F_\theta^+, F_\theta^-)$:

$$\begin{aligned}
\frac{\mathrm{Var}_\theta\big(D_2^{\mathrm{MVD}}(k,m)\big)}{\mathrm{Var}_\theta\big(\tilde{D}_2^{\mathrm{MVD}}(k,m)\big)} &= \frac{2\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right) - \left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right)^2}{2\left(\tilde{F}_\theta^+(q_\alpha(\theta)) + \tilde{F}_\theta^-(q_\alpha(\theta))\right) - \left(\tilde{F}_\theta^+(q_\alpha(\theta)) + \tilde{F}_\theta^-(q_\alpha(\theta))\right)^2} \\
&\quad \cdot \left(\frac{F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))}{\tilde{F}_\theta^+(q_\alpha(\theta)) + \tilde{F}_\theta^-(q_\alpha(\theta))}\right)\left(\frac{\bar{F}_\theta^+(q_\alpha(\theta)) + \bar{F}_\theta^-(q_\alpha(\theta))}{\bar{\tilde{F}}_\theta^+(q_\alpha(\theta)) + \bar{\tilde{F}}_\theta^-(q_\alpha(\theta))}\right) \\
&= \left(1 + \frac{F_\nu^+(q_\alpha(\theta)) + F_\nu^-(q_\alpha(\theta))}{\tilde{F}_\theta^+(q_\alpha(\theta)) + \tilde{F}_\theta^-(q_\alpha(\theta))}\right)\left(1 + \frac{\bar{F}_\nu^+(q_\alpha(\theta)) + \bar{F}_\nu^-(q_\alpha(\theta))}{\bar{\tilde{F}}_\theta^+(q_\alpha(\theta)) + \bar{\tilde{F}}_\theta^-(q_\alpha(\theta))}\right) > 1
\end{aligned}$$

where, for instance, $\bar{F}_\theta^+$ denotes again the complementary distribution function.

3.2.6 Importance Sampling

As the standard derivative estimator is unaffected by the choice of (Z^+, Z^-), despite being the more precise estimator, we need to investigate the efficacy of alternative variance reduction methods. Antithetic variates and stratified sampling are procedures that increase the computational efficiency of measure-valued derivative generations and thus are useful incorporations, see [38], for estimating MVD ranked-data statistics. However, either improvement is incremental as both methods would estimate more expediently the probabilities in Equation (3.5). In addition, the control variate method emphasizes on the choice of map, which is best decided on the specific application and needs to be analytically tractable, rather than the random variable and is thus difficult to analyze. Hence, this variance reduction method will not be considered.

Instead, we analyze the usefulness of the importance sampling method. This method reduces the variance of an estimator by focusing on the random variable generation. For our purposes we will constrain the measure-valued derivative random variables Z^+, Z^-, respectively, within the interval $(Z_{\lceil\alpha m\rceil-1:m-1}, Z_{\lceil\alpha m\rceil:m-1}]$, ensuring that these samples constitute the order statistics $Z^+_{\lceil\alpha m\rceil:m}$, $Z^-_{\lceil\alpha m\rceil:m}$. Specifically, we denote a random variable $\check{Z}^+$ that has support solely within the interval $(Z_{\lceil\alpha m\rceil-1:m-1} - Z_{\lceil\alpha m\rceil:m-1}]$, in which both order statistics are formed from the sequence $\mathbf{Z}\backslash Z_m$, and within this interval $\check{Z}^+ = Z^+$ in distribution. The random variable $\check{Z}^-$ is defined similarly. Let $\check{Z}^+_m$ be an independent copy of $\check{Z}^+$ and be independent of the sequence $\mathbf{Z}\backslash Z_m$. The associated order statistic $\check{Z}^+_{\lceil\alpha m\rceil:m}$ is then the $\lceil\alpha m\rceil^{th}$ smallest realization formed from the collection $\mathbf{Z}\backslash Z_m$ and $\check{Z}^+_m$, and conditioned on $\mathbf{Z}\backslash Z_m$:

$$\begin{aligned}\check{Z}^+_{\lceil\alpha m\rceil:m}\,|\,\mathbf{Z}\backslash Z_m &= \mathrm{ord}_{\alpha,m}\big(Z_1,\ldots,Z_{m-1},\check{Z}^+_m\big) \\ &= \check{Z}^+_m.\end{aligned} \tag{3.13}$$

This last observation occurs by construction. The definition of the order statistic $\check{Z}^-_{\lceil\alpha m\rceil:m}$ similarly follows.

The importance sampling estimator is comparable in thought to the IPA derivative estimator where, in Equation (2.10) of the strong consistency proof, this estimator is construed as the differentiation of a random variable that is restricted within $(Y_1, Y_2]$. In this thought, the importance sampling estimator is the MVD analogue. However, we attain a negative result, as both derivative estimators asymptotically (as $m \to \infty$) equal zero almost surely. Consequently, these proposed importance sampling derivative estimators are both biased and uninformed, unable to ascertain the difference between the value provided by an estimator and the actual value.

From Pyke, [86], the reason for this outcome is that the distributional limit of the spacing, $Z_{\lceil \alpha m\rceil:m-1} - Z_{\lceil \alpha m\rceil-1:m-1}$ is $O(m^{-1})$. As $Z^+ - Z^-$, taking both positive and negative values, is always within a region of $O(m^{-1})$, and by obtaining both these positive and negative values, this difference in expectation is $O(m^{-2})$. To procure this negative outcome we need two additional assumptions for the measure-valued derivative density functions f_θ^+ and f_θ^-:

(B7) In a neighbourhood of $q_\alpha(\theta)$ it holds that both $f_\theta^+(x) > 0$ and $f_\theta^-(x) > 0$.

(B8) Both densities f_θ^+ and f_θ^- are differentiable w.r.t. x with bounded derivative in the neighbourhood of $x = q_\alpha(\theta)$.

We denote for the theorem $\check{D}_l^{\mathrm{MVD}}$, $l = 1,2$ to respectively represent the importance sampling versions of the standard and symmetric derivative estimators.

Theorem 3.2 (Strong Consistency (Importance Sampling))**.** *Suppose that Assumptions* **(B1)**, **(B2)**, *and* **(B4)** *to* **(B8)** *are satisfied. Then under dependence structure* **(B3(a))**, *for $l = 1,2$, it holds almost surely that*

$$\lim_{(m,k)\to(\infty,\infty)} \check{D}_l^{\mathrm{MVD}}(m,k) = 0.$$

As a result the importance sampling versions of both derivative estimators are biased and uninformed.

The derivation of this result is provided in Section 3.4.3.

3.3 Analysis

The components that form the MVD ranked-data derivative estimators, especially the order statistics $Z^+_{\lceil \alpha m\rceil:m}$ and $Z^-_{\lceil \alpha m\rceil:m}$, $\alpha \in (0,1)$, defined in Equation (3.4), are effectively discrete observations with two outcomes: $Z_{\lceil \alpha m\rceil-1:m-1}$ and $Z_{\lceil \alpha m\rceil:m-1}$. From Equation (3.5), each outcome of the order statistic $Z^+_{\lceil \alpha m\rceil:m}$, $Z^-_{\lceil \alpha m\rceil:m}$, respectively, occurs with a probability that depends on the distribution function of each measure-valued derivative random variable. More importantly, letting $\epsilon = O(m^{-1})$, the sum of these two outcomes for either measure-valued derivative order statistic is $1-\epsilon$. As a result, the realizations for the column estimators, $D_{l,j}(m)$, $l = 1,2$, $1 \le j \le k$, is either the spacing term, $Z_{\lceil \alpha m\rceil:m-1} - Z_{\lceil \alpha m\rceil-1:m-1}$, the negative of this value, or zero, and an estimate from either estimator is almost completely the sample average of these outcomes. The precision of the MVD derivative estimators is then the ability to estimate the spacing term. If we compel the realizations of $Z^+_{\lceil \alpha m \rceil:m}$ and $Z^-_{\lceil \alpha m \rceil:m}$ to be Z^+, Z^- within $(Z_{\lceil \alpha m \rceil:m-1}, Z_{\lceil \alpha m \rceil:m-1}]$ respectively, as seen in Section 3.2.6, the measure-valued derivative random variable provides asymptotically no contribution to the estimate for either estimator.

In Chapter 2, the purpose of estimating $Z^+_{\lceil \alpha m \rceil:m}$, $Z^-_{\lceil \alpha m \rceil:m}$, for quantile sensitivity estimation is then a indirect method to estimate $F^+_\theta(q_\alpha(\theta))$, $F^-_\theta(q_\alpha(\theta))$. In particular, the random variable (Z^+, Z^-) does not control the value of the spacing term, but only reduce the variability of these probabilities for the measure-valued derivative order statistics. Even if we only consider this secondary purpose, the more precise estimator is the standard estimator, i.e., dependence between (Z^+, Z^-) within a random variable pair does not improve the estimation of both distribution functions. Mathematically, the reason is that the column estimators are conditionally independent, Equation (3.11). Intuitively, the averaging that occurs within the column estimator, and the changing size of the spacing term that occurs as a result, rendered the dependence between (Z^+, Z^-) meaningless.

For the symmetric estimator, see Chapter 2, positive correlation between (Z^+, Z^-) actually does reduce the variance. This is due to the spacing term for each column estimator remaining fixed and due to the fact that the positive correlation of the random variable pair reduces the variability of outcomes. The choice of random variables $(\tilde{Z}^+, \tilde{Z}^-)$ from the Hahn-Jordan decomposition also has the same effect for this derivative estimator.

Conclusion

This theoretical analysis is in two parts: (i) a presentation of the main results that allows us to infer the behaviour of ranked-data estimation within MVD, and (ii) the proofs, determined in the next section, that gives us the main results. We have ascertained that the principal aspects for either derivative estimator is that column estimates are essentially a discrete random variable predicated on two order statistics, and that the column estimate is greatly influenced by the corresponding spacing. The ability to correlate or choose measure-valued derivative random variables is muted as these random variables do not effect the size of the spacing term. *This is the key reason.* The proofs that provide the main results are constructive, albeit long, and based on minimal, intuitive quantile estimation assumptions.

3.4 Derivations

3.4.1 Asymptotic Variance: Standard Estimator

We begin by taking advantage of the i.i.d. property of the collection of sequences $(\mathbf{Z}(j) : 1 \leq j \leq k)$ the derivative estimator is composed of and the identical distribution of the elements of the random variables constituting order statistics being substituted by their measure-valued derivative counterpart. From Equations (3.1), (3.4), and (3.10):

$$\mathrm{Var}_\theta\big(D_l^{\mathrm{MVD}}(m,k)\big) = \frac{1}{k}\mathrm{Var}_\theta\big(D_{1,1}(m)\big)$$
$$= \frac{c_\theta^2}{k}\Big(m\mathrm{Var}_\theta\big(B_1^+(m)-B_1^-(m)\big) + m(m-1)\mathrm{Cov}_\theta\big(B_1^+(m)-B_1^-(m), B_2^+(m)-B_2^-(m)\big)\Big).$$

To remind, the random variables $B_1^+(m)$, $B_1^-(m)$ are shorthand notation for $Z^+_{\lceil\alpha m\rceil:m}$, $Z^-_{\lceil\alpha m\rceil:m}$, when the random variable Z_m in the sequence $\mathbf{Z}$ is replaced by Z_m^+, Z_m^-. The analogous definition holds for $B_2^\pm(m)$, where the relegated random variable is Z_{m-1}. The leading component of the variance term, as observed in the derivations in Sections 3.4.2, is quadratic in terms of the spacing. From Lemma A.2, part (ii), this implies $m\mathrm{Var}_\theta(B_1^+(m)-B_1^-(m)) \overset{a.s.}{\to} 0$, as $m\to\infty$. In addition, it will be observed that the leading component in terms of different spacing terms in the ensuing Taylor expansion has a combined exponent of two. Therefore, we obtain for the asymptotic variance of the standard estimator:

$$\lim_{m\to\infty}\mathrm{Var}_\theta\left(D_1^{\mathrm{MVD}}(m,k)\right) = \lim_{m\to\infty} m^2\mathrm{Cov}_\theta\big(B_1^+(m)-B_1^-(m), B_2^+(m)-B_2^-(m)\big). \tag{3.14}$$

We continue by modifying the covariance by employing the tower property of conditional expectation. We first condition w.r.t. the common random variables present, for instance, in $B_1^+(m)$, $B_2^+(m)$, namely $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$, before proceeding to condition w.r.t. the sequence $\mathbf{Z}$:

$$\begin{aligned}
&\mathrm{Cov}_\theta\left(B_1^+(m)-B_1^-(m),\, B_2^+(m)-B_2^-(m)\right)\\
&\quad= \mathbb{E}_\theta\big[\big(B_1^+(m)-B_1^-(m)\big)\left(B_2^+(m)-B_2^-(m)\right)\big]\\
&\quad\quad - \mathbb{E}_\theta\big[B_1^+(m)-B_1^-(m)\big]\,\mathbb{E}_\theta\big[B_2^+(m)-B_2^-(m)\big]\\
&\quad= \mathbb{E}\big[\mathbb{E}_\theta\big[\big(B_1^+(m)-B_1^-(m)\big)\left(B_2^+(m)-B_2^-(m)\right)\big|\mathbf{Z}\backslash(Z_{m-1},Z_m)\big]\big]\\
&\quad\quad - \mathbb{E}^2\big[\mathbb{E}_\theta\big[B_1^+(m)-B_1^-(m)\big|\mathbf{Z}\backslash Z_m\big]\big],\\
&\quad= \mathbb{E}\big[\mathbb{E}\big[\mathbb{E}_\theta\big[B_1^+(m)-B_1^-(m)\big|\mathbf{Z}\big]\,\mathbb{E}_\theta\big[B_2^+(m)-B_2^-(m)\big|\mathbf{Z}\big]\big|\mathbf{Z}\backslash(Z_{m-1},Z_m)\big]\big]\\
&\quad\quad - \mathbb{E}^2\big[\mathbb{E}_\theta\big[B_1^+(m)-B_1^-(m)\big|\mathbf{Z}\backslash Z_m\big]\big],
\end{aligned} \tag{3.15}$$

in which the last line is identical to Equation (3.11). As noted in Section 3.2.3, for each product of order statistics within the conditional expectation $\mathbb{E}_\theta[(B_1^+(m) - B_1^-(m))(B_2^+(m) - B_2^-(m))|\mathbf{Z}\backslash(Z_{m-1}, Z_m)]$, the unconditioned measure-valued derivative random variables have a different index, i.e., the unconditioned random variables being Z_{m-1}, Z_{m-1}^-, Z_m, and Z_m^+ for the conditional expectation $\mathbb{E}_\theta[B_1^+(m)B_2^-(m)|\mathbf{Z}\backslash(Z_{m-1}, Z_m)]$. As a result, the final line in Equation (3.15) is permitted by independence.

Since Equation (3.15) has modified the covariance to our needs, we first insert Equation (3.6) into the second term of our expression. This conditional expectation, with $Y_2 > Y_1$ almost surely, is bounded above by the spacing $Y_2 - Y_1$, and the consequent expression has been proven to be uniformly integrable by part (i) of Lemma A.2. We note that the pre-factor $c_\theta > 0$ and that $|F_\theta^+(z) - F_\theta^-(z)| \le 1$:

$$\begin{aligned} mc_\theta\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big] &\le c_\theta \int_{Y_1}^{Y_2} \big|F_\theta^+(z) - F_\theta^-(z)\big|\, dz, \\ &\le c_\theta m(Y_2 - Y_1). \end{aligned} \tag{3.16}$$

We note that (3.16) has already been indirectly proven in Equation (2.16) of the strong consistency proof, Theorem 2.2. This proof method has already been derived in the segment leading to Equation (2.23) in Theorem 2.3. Together with the verification of uniform integrability, the use of the Taylor expansion forms the proof method for attaining the integral expression such as (3.16). We repeat this argument for convenience.

Let $\xi_1 \in (Y_1, Y_2) \subset B_{2r_{m-1}}(q_\alpha(\theta))$, with $B_{2r_{m-1}}(q_\alpha(\theta))$ being the open interval with radius $2r_{m-1}$ defined in Lemma A.2. A two term Taylor expansion of the integrand in (3.16) centred at $z = Y_1$, using a Lagrangian form of the remainder, will attain an upper bound of

$$\begin{aligned} &c_\theta m\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big] \\ &\quad = -\int_{Y_1}^{Y_2} c_\theta m\big(F_\theta^+(Y_1) - F_\theta^-(Y_1)\big) + \big(f_\theta^+(\xi_1) - f_\theta^-(\xi_1)\big)(z - Y_1)\, dz, \\ &\quad \le -c_\theta m\big(F_\theta^+(Y_1) - F_\theta^-(Y_1)\big) + c_\theta c_1 m(Y_2 - Y_1)^2, \end{aligned}$$

where $c_1 \ge |f_\theta^+(y) + f_\theta^-(y)|/2$ for all $y \in B_{2r_{m-1}}(q_\alpha(\theta))$, which is bounded from Assumption **(B6)**. The equivalent lower bound to Equation (3.16) is attained when c_1 is replaced by $-c_1$.

All of the following results are as $m \to \infty$. From [4], Y_1, $Y_2 \overset{a.s.}{\to} q_\alpha(\theta)$, and since F_θ^+, F_θ^- is continuous at $z = q_\alpha(\theta)$, then by the Continuity Mapping Theorem, [101], $F_\theta^+(Y_1) - F_\theta^-(Y_1) \overset{a.s.}{\to} F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))$. As Y_1, Y_2 are formed from i.i.d.

random variables, it follows from [86] that $m(Y_2 - Y_1) \xrightarrow{d} E/f_\theta(q_\alpha(\theta))$, where E is an exponential random variable with mean one, and from Lemma A.2, part (ii), $m(Y_2 - Y_1) \xrightarrow{a.s.} 0$. Connecting these results together with Slutsky's Theorem [45] and the Sandwich Theorem leads us to the distributional result, as $m \to \infty$,

$$c_\theta m \mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big] \xrightarrow{d} -\frac{c_\theta\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right)}{f_\theta(q_\alpha(\theta))}E. \tag{3.17}$$

We note that the RHS in the above equation can be further simplified to the quantile sensitivity expression (for continuous random variables) via the MVD representation of the parameter derivative. However, this is not beneficial to our needs.

Since Equation (3.6) is uniformly integrable via (3.16), interchange of limit and expectation is allowed from the Dominated Convergence Theorem. Together with the Continuity Mapping Theorem, with function $a(x) = x^2$

$$\begin{aligned}
&\lim_{m\to\infty} \mathbb{E}^2\big[c_\theta m \mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big]\big] \\
&\quad = \mathbb{E}^2\Big[\lim_{m\to\infty} c_\theta m \mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big]\Big] \\
&\quad = -\frac{c_\theta^2\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right)^2}{f_\theta^2(q_\alpha(\theta))}.
\end{aligned} \tag{3.18}$$

We now proceed to the second and major component of the proof. Let $Y_1^{(1)} = Y_1$, $Y_2^{(1)} = Y_2$. Similarly, define $Y_1^{(2)}$, $Y_2^{(2)}$ to be equivalent order statistics to Y_1, Y_2 by

$$\begin{aligned}
Y_1^{(2)} &= \mathrm{ord}_{\alpha_1,m}(Z_1,\ldots,Z_{m-2},Z_m), \quad \text{and,} \\
Y_2^{(2)} &= \mathrm{ord}_{\alpha_2,m}(Z_1,\ldots,Z_{m-2},Z_m).
\end{aligned}$$

If we suppose Assumption **(B3(c))**, the other two conditional expectations in Equation (3.15), similar in representation to Equation (3.8) and written in (3.19) below, may not have a Lebesgue absolute continuous density. Heeding Section 3.2.2, the functions $F_\theta^+(\cdot|Z_m)$, $F_\theta^+(\cdot|Z_m)$, are regular conditional distributions. As Equation (3.8) was attained via the Stieltjes integration by parts formula, the argument for the regular conditional distributions is in terms of the left limit, for instance,

$$\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\big] = -\int_{(Y_1^{(1)},Y_2^{(1)}]} F_\theta^+(z^-|Z_m) - F_\theta^-(z^-|Z_m)\,dz.$$

The integration is w.r.t. a semi-open interval due to the definition of outcomes given in (3.5), and this is equal with respect to the equivalent open interval where the regular conditional distribution are written w.r.t. the standard argument:

$$
\begin{aligned}
\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\,\big|\,\mathbf{Z}\big] &= -\int_{(Y_1^{(1)},Y_2^{(1)})} F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\,dz \\
\mathbb{E}_\theta\big[B_2^+(m) - B_2^-(m)\,\big|\,\mathbf{Z}\big] &= -\int_{(Y_1^{(2)},Y_2^{(2)})} F_\theta^+(z|Z_{m-1}) - F_\theta^-(z|Z_{m-1})\,dz. \qquad (3.19)
\end{aligned}
$$

To progress further, we need to separate the random variables Z_{m-1} and Z_m, from the order statistics $Y_1^{(l)}$, $Y_2^{(l)}$, $l = 1,2$, in (3.19). This avoids difficulties that would otherwise occur when the conditional expectation w.r.t. $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$ is computed for the product of these terms. To do this, we use the same argument that culminates to Equation (3.5). We insert respectively Z_{m-1} and Z_m, into the collection of random variables $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$ to attain equivalent outcomes. Let $W_1 = Z_{\lceil\alpha m\rceil-2:m-2}$, $W_2 = Z_{\lceil\alpha m\rceil-1:m-2}$, and $W_3 = Z_{\lceil\alpha m\rceil:m-2}$ be order statistics consisting of $m-2$ elements formed from $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$. For the order statistics $Y_1^{(1)}$, the three possible outcomes are W_1, W_2, Z_{m-1}, depending on the realization Z_{m-1}, and for $Y_1^{(2)}$, either W_2, W_3, or Z_{m-1}. The relationship between $Y_1^{(l)}$, $Y_2^{(l)}$, $l = 1,2$, and their outcomes occur with probability one and are given below:

$$
\begin{aligned}
Y_1^{(1)} &= W_1\,1\{Z_{m-1} \le W_1\} + Z_{m-1}\,1\{Z_{m-1} \in (W_1, W_2]\} + W_2\,1\{Z_{m-1} > Y_2\}, \quad \text{and} \\
Y_1^{(2)} &= W_2\,1\{Z_{m-1} \le W_2\} + Z_m\,1\{Z_{m-1} \in (W_2, W_3]\} + W_3\,1\{Z_{m-1} > Y_3\}.
\end{aligned}
$$

By pairing these relations together we attain the needed results

$$
\left(Y_1^{(1)}, Y_2^{(1)}\right)\Big|\,\mathbf{Z} = \begin{cases}
\textbf{Random Variable} & \textbf{Event} \\
(W_1, W_2) & \{\omega : Z_{m-1} \le W_1\} \\
(Z_{m-1}, W_2) & \{\omega : W_1 < Z_{m-1} \le W_2\} \\
(W_2, Z_{m-1}) & \{\omega : W_2 < Z_{m-1} \le W_3\} \\
(W_2, W_3) & \{\omega : Z_{m-1} > W_3\}.
\end{cases}
$$

A similar almost sure relation is acquired for the pair $(Y_1^{(2)}, Y_2^{(2)})$, by substituting Z_{m-1} with Z_m. Via the Continuous Mapping Theorem, since an integral is continuous w.r.t. its terminals, incorporating this relation into $\mathbb{E}_\theta[B_1^+(m) - B_1^-(m)|\mathbf{Z}]$ we subsequently have four integrals, with caution justified due to the effect thatZ_m has on the conditional expectation w.r.t. $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$

$$
\begin{aligned}
&\mathbb{E}_\theta\big[B_1^+(m) - B_1^-(m)\big|\mathbf{Z}\big] \\
&\quad = \underbrace{-\int_{(W_1,W_2)} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{Z_{m-1} \le W_1\}\, dz}_{(1a)} \\
&\quad\quad \underbrace{-\int_{(Z_{m-1},W_2)} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{W_1 < Z_{m-1} \le W_2\}\, dz}_{(1b)} \\
&\quad\quad \underbrace{-\int_{(W_2,Z_{m-1})} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{W_2 < Z_{m-1} \le W_3\}\, dz}_{(1b)} \\
&\quad\quad \underbrace{-\int_{(W_2,W_3)} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{Z_{m-1} > W_3\}\, dz}_{(1a)}. \qquad (3.20)
\end{aligned}
$$

The labels $(1a)$ and $(1b)$ indicate the different leading exponents w.r.t. a monomial of a spacing. For the conditional expectation $\mathbb{E}_\theta[B_2^+(m) - B_2^-(m)|\mathbf{Z}]$, the difference in expression to (3.20) is the interchange of random variables Z_{m-1} and Z_m, supplanting both the σ-field induced by the random variable for the regular conditional distributions and the integral terminal bounds. These integrals are labelled $(2a)$ and $(2b)$ for the same above reason.

From the sixteen terms that arise from the product of these conditional expectations, only four of these terms will attain an almost sure non-zero contribution as $m \to \infty$. These are the products that occur from $(1a)\times(2a)$. The conditional expectations conditioned on the sequence $\mathbf{Z}\backslash(Z_{m-1}, Z_m)$ of these leading terms are given below:

$$
\begin{aligned}
&\mathbb{E}_\theta\big[\big(B_1^+(m) - B_1^-(m)\big)\big(B_2^+(m) - B_2^-(m)\big)\big|\mathbf{Z}\backslash(Z_{m-1}, Z_m)\big] \\
&= \mathbb{E}_\theta\Bigg[\left(\int_{(W_1,W_2)} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{Z_{m-1} \le W_1\}\, dz\right) \\
&\quad \cdot \left(\int_{(W_1,W_2)} \big(F_\theta^+(w|Z_{m-1}) - F_\theta^-(w|Z_{m-1})\big)\, 1\{Z_m \le W_1\}\, dw\right)\Bigg|\mathbf{Z}\backslash(Z_{m-1}, Z_m)\Bigg] \\
&\quad + 2\mathbb{E}_\theta\Bigg[\left(\int_{(W_1,W_2)} \big(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\big)\, 1\{Z_{m-1} \le W_1\}\, dz\right) \\
&\quad \cdot \left(\int_{(W_2,W_3)} \big(F_\theta^+(w|Z_{m-1}) - F_\theta^-(w|Z_{m-1})\big)\, 1\{Z_m > W_3\}\, dw\right)\Bigg|\mathbf{Z}\backslash(Z_{m-1}, Z_m)\Bigg]
\end{aligned}
$$

$$+\,\mathbb{E}_\theta\left[\left(\int_{(W_2,W_3)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_{m-1}>W_3\}\,dz\right)\right.$$
$$\left.\cdot\left(\int_{(W_2,W_3)}\left(F_\theta^+(w|Z_{m-1})-F_\theta^-(w|Z_{m-1})\right)1\{Z_m>W_3\}\,dw\right)\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right].$$

For the middle term, a simplification was utilized due to an identity in distribution. As the random variables Z_{m-1} and Z_m are independent of each other, the above expression is further simplified. The simplification is only tantamount to a second moment of a conditional expectation:

$$=\mathbb{E}^2\left[\int_{(W_1,W_2)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m\le W_1\}\,dz\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]$$
$$+\,\mathbb{E}^2\left[\int_{(W_2,W_3)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m>W_3\}\,dz\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]$$
$$+\,2\mathbb{E}_\theta\left[\int_{(W_1,W_2)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m>W_3\}\,dz\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]$$
$$\cdot\,\mathbb{E}_\theta\left[\int_{(W_2,W_3)}\left(F_\theta^+(w|Z_{m-1})-F_\theta^-(w|Z_{m-1})\right)1\{Z_{m-1}\le W_1\}\,dw\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right].\tag{3.21}$$

We denote by G_θ^+, G_θ^- the distribution functions for the respective random variable pairs (Z^+,Z), (Z^-,Z). To compute these conditional expressions, we repeat the argument that preceded and ended with Equation (3.9). We make use of the Disintegration Theorem, [62], and Fubini's Theorem to attain the associate bivariate Stieltjes measures via, for instance $dG_\theta^+(u,v)=dF_\theta^+(u|v)\,dF_\theta(v)$, before converting the integral into one or two distribution functions. We begin at the top expression and continue downward.

$$\mathbb{E}\left[\int_{(W_1,W_2)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m\le W_1\}\,dz\middle|\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]$$
$$=\int_{(-\infty,W_1]}\int_{(W_1,W_2)}\left(\int_{(-\infty,z]}\left(dF_\theta^+(u|v)-dF_\theta^-(u|v)\right)\right)dz\,dF_\theta(v)$$
$$=\int_{(W_1,W_2)}\int_{(-\infty,z]\times(-\infty,W_1]}\left(dG_\theta^+(u,v)-dG_\theta^-(u,v)\right)dz$$
$$=\int_{(W_1,W_2)}G_\theta^+(z,W_1)\,-\,G_\theta^-(z,W_1)\,dz.\tag{3.22}$$

Let $B^2_{2r_m}(q_\alpha(\theta))$, defined below, be the open neighbourhood centred at $(q_\alpha(\theta), q_\alpha(\theta))$ with radius $2r_{m-2}$ defined from the metric induced by the sup-norm. The term r_m is the upper bound in Equation (A.1) in Lemma A.1:

$$B^2_{2r_m}(q_\alpha(\theta)) = \left\{(x_1, x_2) : \|(x_1, x_2) - (q_\alpha(\theta), q_\alpha(\theta))\|_\infty < \frac{2}{f_\theta(q_\alpha(\theta))}\frac{(\ln(m))^{\frac{1}{2}}}{m^{\frac{1}{2}}}\right\}.$$

The distribution functions G^+_θ, G^-_θ are continuous on $B^2_{2r_{m-2}}(q_\alpha(\theta))$, containing the order statistics W_l, $l = 1,2,3$. This statement holds as F_θ, F^+_θ, F^-_θ are each continuous on $B_{2r_{m-2}}(q_\alpha(\theta))$. The derivation is given in Postscript I after the main proof. The purpose of the preceding definition and proof is that we require neighbourhood continuity for a Stieltjes measure to behave as standard Lebesgue measure. In general, a Lebesgue-Stieltjes does not possess this behaviour and we have to consider the behaviour of the Stieltjes measure, and the region of integration when employing the Stieltjes integral. Let $\xi_2 \in (W_1, W_2) \subset B_{2r_{m-2}}(q_\alpha(\theta))$. Given that G^+_θ, G^-_θ are continuous, we conclude then from the Mean-Value Theorem for Integration, [78]:

$$\begin{aligned}\mathbb{E}&\left[m\int_{(W_1,W_2]}\left(F^+_\theta(z|Z_m) - F^-_\theta(z|Z_m)\right)1\{Z_m \le W_1\}\,dz\,\middle|\,\mathbf{Z}\backslash(Z_{m-1}, Z_m)\right]\\ &= m\left(G^+_\theta(\xi_2, W_1) - G^-_\theta(\xi_2, W_1)\right)(W_2 - W_1).\end{aligned}$$

Since $|G^+_\theta|$, $|G^-_\theta| \le 1$, this expression has finite expectation for any m, which follows from part (i) of Lemma A.2.

To attain the limit, as $m \to \infty$, that led to Equation (3.17), we implement the same collection of limiting arguments. The only variation is due to the bivariate distribution functions. As $W_1 < \xi_2 < W_2$, and both $W_1, W_2 \overset{a.s.}{\to} q_\alpha(\theta)$, $\xi_2 \overset{a.s.}{\to} q_\alpha(\theta)$, and therefore, via the Continuous Mapping Theorem, $G^+_\theta(\xi_2, W_1) \overset{a.s.}{\to} G^+_\theta(q_\alpha(\theta), q_\alpha(\theta))$ with the same result for the negative variant. Then from Slutsky's Theorem, in distribution

$$\begin{aligned}\lim_{m\to\infty}\mathbb{E}&\left[\int_{(W_1,W_2]}\left(F^+_\theta(z|Z_m) - F^-_\theta(z|Z_m)\right)1\{Z_m \le W_1\}\,dz\,\middle|\,\mathbf{Z}\backslash(Z_{m-1}, Z_m)\right]\\ &= \frac{G^+_\theta(q_\alpha(\theta), q_\alpha(\theta)) - G^-_\theta(q_\alpha(\theta), q_\alpha(\theta))}{f_\theta(q_\alpha(\theta))}E_1,\end{aligned} \tag{3.23}$$

where E_1 is an exponential mean one random variable from the spacing limit.

The other expressions in (3.21) are attained in the same manner:

$$\begin{aligned}
&\mathbb{E}\left[\int_{(W_2,W_3)} m\left(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\right) 1\{Z_m > W_3\}\, dz \,\middle|\, \mathbf{Z}\backslash(Z_{m-1}, Z_m)\right] \\
&= \int_{(W_3,\infty)} \int_{(W_2,W_3)} \int_{(-\infty,z]} m\left(dF_\theta^+(u|v) - dF_\theta^-(u|v)\right) dz\, dF_\theta(v) \\
&= \int_{(W_2,W_3)} \int_{(-\infty,z]\times(W_3,\infty)} m\left(dG_\theta^+(u,v) - dG_\theta^-(u,v)\right) dz \\
&= \int_{(W_2,W_3)} m\left(F_\theta^+(z) - F_\theta^-(z)\right) - \left(G_\theta^+(z,W_3) - G_\theta^-(z,W_3)\right) dz,
\end{aligned}$$

via continuity for the last operation. Let $\xi_3 \in (W_2, W_3)$. By the Mean-Value Theorem for Integration, the above equation equals

$$= m\left(\left(F_\theta^+(\xi_3) - F_\theta^-(\xi_3)\right) - \left(G_\theta^+(\xi_3, W_3) - G_\theta^-(\xi_3, W_3)\right)\right)(W_3 - W_2),$$

and as $m \to \infty$

$$\xrightarrow{d} \frac{\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right) - \left(G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) - G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))\right)}{f_\theta(q_\alpha(\theta))} E_2. \quad (3.24)$$

The same spacing limit arose from a different sequence of spacing terms. From Pyke, [86], if the two spacing sequences do not stem from adjacent order statistics, then the exponential random variables E_1, E_2, in the limit are independent. In Postscript II, we show that independence between E_1, E_2, also occurs when the order statistics are adjacent. Similarly to Equations (3.23) and (3.24), as $m \to \infty$

$$\begin{aligned}
&\mathbb{E}_\theta\left[\int_{(W_1,W_2)} \left(F_\theta^+(z|Z_m) - F_\theta^-(z|Z_m)\right) 1\{Z_m > W_3\}\, dz \,\middle|\, \mathbf{Z}\backslash(Z_{m-1}, Z_m)\right] \\
&\xrightarrow{d} \frac{\left(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta))\right) - \left(G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) - G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))\right)}{f_\theta(q_\alpha(\theta))} E_1, \quad \text{and} \\
&\mathbb{E}_\theta\left[\int_{(W_2,W_3)} \left(F_\theta^+(w|Z_{m-1}) - F_\theta^-(w|Z_{m-1})\right) 1\{Z_{m-1} \le W_1\}\, dw \,\middle|\, \mathbf{Z}\backslash(Z_{m-1}, Z_m)\right] \\
&\xrightarrow{d} \frac{G_\theta^+(q_\alpha(\theta), q_\alpha(\theta)) - G_\theta^-(q_\alpha(\theta), q_\alpha(\theta))}{f_\theta(q_\alpha(\theta))} E_2. \quad (3.25)
\end{aligned}$$

The product of integrals that are not labelled $(1a) \times (2a)$, in Equation (3.20) and immediately thereafter, are shown to be almost surely zero via related operations. We present two absolute bounds, one from $(1a) \times (2b)$ and the other from

$(2a)\times(2b)$, and claim that the method of solution and form of solution is the same for each class of expression. We begin with the combination $(1a)\times(2b)$:

$$\Bigg|\mathbb{E}_\theta\Bigg[m^2\Bigg(\int_{(W_1,W_2)}\big(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\big)\,1\{Z_{m-1}\le W_1\}\,dz\Bigg)$$
$$\cdot\Bigg(\int_{(W_2,Z_m)}\big(F_\theta^+(w|Z_{m-1})-F_\theta^-(w|Z_{m-1})\big)\,1\{W_2<Z_m\le W_3\}\,dw\Bigg)\Bigg|\mathbf{Z}\backslash(Z_{m-1},Z_m)\Bigg]\Bigg|$$
$$\le\mathbb{E}_\theta\big[m^2(W_2-W_1)(Z_m-W_2)\,1\{Z_{m-1}\le W_1\}\,1\{W_2\le Z_m\le W_3\}\,\big|\mathbf{Z}\backslash(Z_{m-1},Z_m)\big]$$

since $|F_\theta^+(\cdot|\cdot)-F_\theta^-(\cdot|\cdot)|\le 1$. Then by independence

$$\le m^2(W_2-W_1)(W_3-W_2)F_\theta(W_1)\,(F_\theta(W_3)-F_\theta(W_2))$$
$$= m^2(W_2-W_1)(W_3-W_2)^2F_\theta(W_1)\,f_\theta(\xi_4),$$

due to the Mean-Value Theorem and $\xi_4\in(W_2,W_3)\subset B_{2r_{m-2}}(q_\alpha(\theta))$. From Lemma A.2, especially part (ii) where $m(W_3-W_2)^2\overset{a.s.}{\to}0$, Slutsky's Theorem, multiplying the limiting results together, and [96], property H, p. 185, the absolute bound almost surely converges to zero as $m\to\infty$. Administering the same arguments as before, for combination $(1b)\times(2b)$:

$$\Bigg|\mathbb{E}_\theta\Bigg[m^2\Bigg(\int_{(Z_{m-1},W_2)}\big(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\big)\,1\{W_1<Z_{m-1}\le W_2\}\,dz\Bigg)$$
$$\cdot\Bigg(\int_{(W_2,Z_m)}\big(F_\theta^+(w|Z_{m-1})-F_\theta^-(w|Z_{m-1})\big)\,1\{W_2<Z_m\le W_3\}\,dw\Bigg)\Bigg|\mathbf{Z}\backslash(Z_{m-1},Z_m)\Bigg]\Bigg|$$
$$\le\mathbb{E}_\theta\big[m^2(W_2-Z_{m-1})(Z_m-W_2)1\{W_1<Z_{m-1}\le W_2\}1\{W_2<Z_m\le W_3\}\big|\mathbf{Z}\backslash(Z_{m-1},Z_m)\big],$$

and via independence,

$$\le m^2(W_2-W_1)(W_3-W_2)\,(F_\theta(W_2)-F_\theta(W_1))\,(F_\theta(W_3)-F_\theta(W_2))$$
$$= m^2(W_2-W_1)^2(W_3-W_2)^2f_\theta(\xi_4)f_\theta(\xi_5),$$

in which $\xi_5\in(W_1,W_2)\subset B_{2r_{m-2}}(q_\alpha(\theta))$. As both $m(W_2-W_1)^2\overset{a.s.}{\to}0$, $m(W_3-W_2)^2\overset{a.s.}{\to}0$ as $m\to\infty$, Lemma A.2, the absolute bound almost surely converges to zero.

As each of the expressions, (3.23) - (3.25), are uniformly integrable, the Dominated Convergence Theorem can be applied to (3.21). Combined with the mappings $a(x)=x^2$ and $b(x,y)=xy$ where needed, the Continuous Mapping Theorem yields the limiting distributional result as $m\to\infty$ for Equation (3.15), which can now be written as a second moment:

$$
\begin{aligned}
&\lim_{m\to\infty} \mathbb{E}\big[\,\mathbb{E}_\theta\big[m^2\left(B_1^+(m)-B_1^-(m)\right)\left(B_2^+(m)-B_2^-(m)\right)\big|\,\mathbf{Z}\backslash(Z_{m-1},Z_m)\big]\big]\\
&= \mathbb{E}\left[\lim_{m\to\infty}\mathbb{E}_\theta^2\left[m^2\int_{(W_1,W_2)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m\le W_1\}\,dz\,\middle|\,\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]\right.\\
&\quad+\lim_{m\to\infty}\mathbb{E}_\theta^2\left[m^2\int_{(W_2,W_3)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m> W_3\}\,dz\,\middle|\,\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]\\
&\quad+2\lim_{m\to\infty}\mathbb{E}_\theta\left[m\int_{(W_1,W_2)}\left(F_\theta^+(z|Z_m)-F_\theta^-(z|Z_m)\right)1\{Z_m> W_3\}\,dz\,\middle|\,\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]\\
&\quad\left.\cdot\,\mathbb{E}_\theta\left[m\int_{(W_2,W_3)}\left(F_\theta^+(w|Z_{m-1})-F_\theta^-(w|Z_{m-1})\right)1\{Z_{m-1}\le W_1\}\,dw\,\middle|\,\mathbf{Z}\backslash(Z_{m-1},Z_m)\right]\right]\\
&=\frac{1}{f_\theta^2(q_\alpha(\theta))}\,\mathbb{E}_\theta\Big[\Big(\left(F_\theta^+(q_\alpha(\theta))-F_\theta^-(q_\alpha(\theta))\right)E_2\\
&\qquad\qquad+\left(G_\theta^+(q_\alpha(\theta),q_\alpha(\theta))-G_\theta^-(q_\alpha(\theta),q_\alpha(\theta))\right)(E_1-E_2)\Big)^2\Big].
\end{aligned}
$$

As $\mathbb{E}[E_2^2]=2$, $\mathbb{E}[E_2(E_1-E_2)]=-1$, and $\mathbb{E}[(E_1-E_2)^2]=2$ by independence, this previous expectation has limit as $m\to\infty$ of

$$
\begin{aligned}
&=\frac{2}{f_\theta^2(q_\alpha(\theta))}\Big(\left(F_\theta^+(q_\alpha(\theta))-F_\theta^-(q_\alpha(\theta))\right)^2+\left(G_\theta^+(q_\alpha(\theta),q_\alpha(\theta))-G_\theta^-(q_\alpha(\theta),q_\alpha(\theta))\right)^2\\
&\qquad+\left(F_\theta^+(q_\alpha(\theta))-F_\theta^-(q_\alpha(\theta))\right)\left(G_\theta^+(q_\alpha(\theta),q_\alpha(\theta))-G_\theta^-(q_\alpha(\theta),q_\alpha(\theta))\right)\Big).
\end{aligned}
$$

This expression together with Equation (3.18) provides us with the components to ascertain the asymptotic variance according to (3.14). After a little algebra, observing Assumption **(B3(c))**, the same expression is attained for the standard derivative estimator (this is part (ii)). □

Postscript I: We only ascertain the continuity of G_θ^+ on $B^2_{2r_{m-2}}(q_\alpha(\theta))$. The derivations for the other bivariate distributions are the same. The definition of continuity we will verify is the ϵ-δ condition presented in Pugh, [85], where both bounds are represented in the same metric. For our purposes, this will be metric induced by the sup-norm. As $G_\theta:\mathbb{R}^2\mapsto\mathbb{R}$, we note that all l^p norms, $p\ge 1$, are identical.

This derivation uses the operator Δ_{a_l,b_l}, for $l=1,2$, used in [12] for instance, to signify the difference of two distribution functions at a given element. The value of the index l indicates which argument the operator is acting on: for example, $\Delta_{a_1,b_1}G_\theta^+(x,y)=G_\theta^+(a_1,y)-G_\theta^+(b_1,y)$. The argument x^- will again represent the left limit towards the specified value.

Let $s_1 < t_1$ and $s_2 < t_2$, such that $(s_1, t_1) \times (s_2, t_2) \subset B^2_{2r_{m-2}}(q_\alpha(\theta))$, with $B^2_{2r_{m-2}}(q_\alpha(\theta))$ an open ball in the l^∞ metric. Then from the provided assumptions of continuity, **(B2)** and **(B6)**, for every $\epsilon > 0$, $\exists \delta_1, \delta_2 > 0$ such that $|F^+_\theta(t_1) - F^+_\theta(s_1)| < \epsilon$ and $|F_\theta(t_2) - F_\theta(s_2)| < \epsilon$. Then for $\delta = \min\{\delta_1, \delta_2\}$:

$$\begin{aligned}\Delta_{s_1,t_1^-}\Delta_{s_2,t_2^-} G^+_\theta(x,y) &= \int_{(s_1,t_1)\times(s_2,t_2)} dG^+_\theta(u,v) \\ &= \frac{1}{2}\left(\int_{(s_1,t_1)}\int_{(s_2,t_2)} dF_\theta(v|u)\, dF^+_\theta(u) + \int_{(s_2,t_2)}\int_{(s_1,t_1)} dF^+_\theta(u|v)\, dF_\theta(v)\right).\end{aligned}$$

This above expression is due to the Disintegration Theorem, [62], and the existence of a regular conditional distribution, [96]. The desired result occurs from the bound of a distribution function:

$$\begin{aligned}&\le \frac{1}{2}\left(\int_{(s_1,t_1)} dF^+_\theta(u) + \int_{(s_2,t_2)} dF_\theta(v)\right) \\ &< \frac{1}{2}\cdot 2\epsilon = \epsilon.\end{aligned}$$

Postscript II: Let $(U_i : 1 \le l \le n)$ be n i.i.d. copies of a standard uniform, $U \in (0,1)$, random variable. Additionally, let X_1 be a random variable such that $\mathbb{E}|X_1| < \infty$ and $(X_i : 1 \le l \le n)$ be an i.i.d. collection. From [91], the finite absolute expectation implies the finiteness of any order statistic $X_{\lceil \alpha n\rceil:n}$, $\alpha \in (0,1)$.

Let $U_{l:n}$, $l = r-1, r, r+1$, be three consecutive standard uniform order statistics, and $r = 2, \ldots, n-1$. We commence this computation from the joint density function of $(n(U_{r:n} - U_{r-1:n}), n(U_{r+1:n} - U_{r:n}))$ following the standard representation of depicting density functions of order statistics, see [18]. We then employ the Mean-Value Theorem to convert the result attained from the uniform distribution to the general case, the bivariate random variable pair $(n(X_{r:n} - X_{r-1:n}), n(X_{r+1:n} - X_{r:n}))$. The required distributional spacing limit and asymptotic independence occurs when we take $n \to \infty$. This result also holds for $r = 1$, and $r = n$, defining $U_{0:n} = 0$ and $U_{n+1:n} = 1$, but we do not present this calculation.

We begin by

$$\begin{aligned}&\mathbb{P}(n(U_{r:n} - U_{r-1:n}) \le z,\, n(U_{r+1:n} - U_{r:n}) \le w) \\ &\quad = \mathbb{P}\Big(U_{r:n} - U_{r-1:n} \le \frac{z}{n},\, U_{r+1:n} - U_{r-1:n} \le \frac{z+w}{n}\Big)\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \mathbb{P}\Big(U_{r:n} \le \frac{z}{n} + u,\, U_{r+1:n} \le \frac{z+w}{n} + u,\, U_{r-1:n} \in du\Big) \\
&= \int_{-\infty}^{\infty} \mathbb{P}\Big(U_{r:n} \le \frac{s}{n} + u,\, U_{r+1:n} \le \frac{t}{n} + u,\, U_{r-1:n} \in du\Big),
\end{aligned}
$$

where $s = z$, $t = w+z$, and $t > s$. Let p denote the density function of the spacing pair $(n(U_{r:n} - U_{r-1:n}), n(U_{r+1:n} - U_{r:n}))$ and q the density function of the order statistic triple $(U_{r-1:n}, U_{r:n}, U_{r+1:n})$. As the standard uniform distribution is continuous with differentiable density on $(0,1)$, the integrand above is differentiable w.r.t. w and z and differentiation under the integral sign is allowed, [66], and

$$
\begin{aligned}
p(z,w) &= \int_{-\infty}^{\infty} \frac{\partial^2}{\partial s \partial t} \mathbb{P}\Big(U_{r:n} \le \frac{s}{n} + u,\, U_{r+1:n} \le \frac{t}{n} + u,\, U_{r-1:n} \in du\Big) \\
&= \int_0^{1-\frac{z+w}{n}} \frac{1}{n^2} q\Big(u, u + \frac{z}{n},\, u\frac{z+w}{n}\Big)\, du \\
&= \frac{1}{n^2} \frac{n!}{(r-2)!(n-r-1))!} \int_0^1 u^{r-2} \Big(1 - u - \frac{z+w}{n}\Big)^{n-r-1}\, du.
\end{aligned}
$$

The order statistic triple $(U_{r-1:n}, U_{r:n}, U_{r+1:n})$ is equivalent to have $r-2$ random variables less than u, $n-(r+1)$ greater than $1-u-(w+z)/n$ and the middle three specified at the values of their arguments. The number of possible combinations for an i.i.d. uniform sample of size n is given by the multinomial coefficient. We now implement the substitution $u = v(1-(z+w)/n)$ to convert the above integral into a beta function before representing the beta function and the factorial terms by the gamma function. This leads us to the term sequence that defines the constant e:

$$
\begin{aligned}
&= \frac{1}{n^2} \frac{\Gamma(n+1)}{\Gamma(r-1)\Gamma(n-r)} \int_0^1 v^{r-2}(1-v)^{n-r-1} \Big(1 - \frac{z+w}{n}\Big)^{n-2}\, dv \\
&= \frac{1}{n^2} \frac{\Gamma(n+1)}{\Gamma(r-1)\Gamma(n-r)} \cdot \frac{\Gamma(r-1)\Gamma(n-r)}{\Gamma(n-1)} \Big(1 - \frac{z+w}{n}\Big)^{n-2} \\
&= \frac{n-1}{n} \Big(1 - \frac{z+w}{n}\Big)^{n-2}.
\end{aligned}
$$

This last term has a limit of $e^{-(z+w)}$ as $n \to \infty$. Therefore, successive normalized standard uniform-distributed spacings are asymptotically independent and both limits in distribution are equal to the exponential distribution with mean one.

For the general case, we let X and U be connected via the inverse distribution function, $X_{r:n} = F^{-1}(U_{r:n})$ where $F \sim X$. The akin density function is denoted by f. By the Mean-Value Theorem, the two spacing terms are related by

$$\begin{aligned} X_{r:n} - X_{r-1:n} &= F^{-1}(U_{r:n}) - F^{-1}(U_{r-1:n}) \\ &= \frac{1}{f\big(F^{-1}(\zeta)\big)}(U_{r:n} - U_{r-1:n}) \end{aligned}$$

with $\zeta \in (U_{r:n} - U_{r-1:n})$. If $r/n \to \alpha \in (0,1)$, as $n \to \infty$, $F^{-1}(\zeta) = q_\alpha$, the quantile function for X. Then by the Continuous Mapping Theorem and Slutsky's Theorem,

$$n(X_{r:n} - X_{r-1:n}) \xrightarrow{d} \frac{E_1}{f(q_\alpha)}$$

where $E_1 \sim Exp(1)$.

3.4.2 Asymptotic Variance: Symmetric Estimator:

This proof continues from Equation (3.12). We denote $Z^+_{\lceil \alpha m \rceil:m}$, $Z^-_{\lceil \alpha m \rceil:m}$ as order statistics in which, successively, the random variable Z_m is substituted for Z^+_m and Z^-_m, copies of random variables Z^+, Z^-. For the calculations, though, we use $B^+(m)$, $B^-(m)$ to condense the notation when representing these order statistics. From the present position the only result that remains to be solved is the limit for the second moment. As $m \to \infty$:

$$\begin{aligned} \lim_{m\to\infty} \mathrm{Var}_\theta(D_2^{\mathrm{MVD}}(m,k)) = \frac{c_\theta^2}{k}\Big(&\lim_{m\to\infty} \mathbb{E}\Big[\mathbb{E}_\theta\Big[m^2\left(B^+(m) - B^-(m)\right)^2 \Big| \mathbf{Z}\backslash Z_m\Big]\Big] \\ &- \lim_{m\to\infty} \mathbb{E}^2\big[\mathbb{E}_\theta\big[m\left(B^+(m) - B^-(m)\right)\big| \mathbf{Z}\backslash(Z_{m-1}, Z_m)\big]\big]\Big). \end{aligned} \tag{3.26}$$

The limiting result for the second term in (3.26) has already been provided in Equation (3.18), i.e.,

$$\lim_{m\to\infty} \mathbb{E}^2\big[\mathbb{E}_\theta\big[m\left(B^+(m) - B^-(m)\right)\big| \mathbf{Z}\backslash(Z_{m-1}, Z_m)\big]\big] = -\frac{\left(F^+_\theta(q_\alpha(\theta)) - F^-_\theta(q_\alpha(\theta))\right)^2}{f^2_\theta(q_\alpha(\theta))}, \tag{3.27}$$

from a combination of Equation (2.21) in the derivation of the strong consistency property of both MVD derivative estimators, Theorem 2.2, and the Continuous Mapping Theorem. A synopsis of this derivation preceded (3.18).

In this proof, we first derive the asymptotic variance for part (i), supposing mutual independence, Assumption **(B3(a))**. Then in part (ii) we repeat the computation supposing Assumption **(B3(c))**. In this case, there is dependence within at least one of the random variable pairs (Z^+, Z), (Z^-, Z), and there may be dependence between the the measure-valued derivative random pair (Z^+, Z^-). In part (ii) of the derivation, we make use of the calculation conducted for the independent case and only concentrate on the additional conditional expectations needed to compute the limiting result. Already introduced in Section 3.4.1, this derivation is conducted through the Stieltjes integral, employing regular conditional distributions to meet the theoretical requirements. The Disintegration Theorem is frequently used in the computation and is key for the completion of this proof. The asymptotic variance derivation supposing Assumption **(B3(c))** also subsumes Assumption **(B3(b))**, where the dependence is not through Z. The same limiting result heeding Assumption **(B3(b))** is attained by the same procedure but without the initial need to condition w.r.t. $\mathbf{Z}$ before implementing the tower property of conditional expectation.

Part (i): Due to the mutual independence, the conditional second moment on the LHS in the below expression reduces to:

$$\mathbb{E}_\theta\Big[\big(B^+(m) - B^-(m)\big)^2\Big|\mathbf{Z}\backslash Z_m\Big] = \mathbb{E}_\theta\Big[\big(B^+(m)\big)^2\Big|\mathbf{Z}\backslash(Z_{m-1}, Z_m)\Big] + \mathbb{E}_\theta\big[(B^-(m))^2\big|\mathbf{Z}\backslash Z_m\big] \\ + 2\mathbb{E}_\theta\big[B^+(m)\big|\mathbf{Z}\backslash Z_m\big]\,\mathbb{E}_\theta[B^-(m)|\mathbf{Z}\backslash Z_m]\,. \tag{3.28}$$

This reduction is a combination of conditional first and second moments of order statistics. The equivalent outcomes for the first moments are presented in Equation (3.5), and the equivalent outcomes for the second moments are argumentatively similar. Let $Y_1 = Z_{\lceil\alpha m\rceil-1:m-1}$, $Y_2 = Z_{\lceil\alpha m\rceil:m-1}$, be the before described order statistics of $m-1$ elements comprised from the sequence $\mathbf{Z}\backslash Z_m$. By the above assertion

$$\mathbb{E}_\theta\Big[\big(B^+(m)\big)^2\Big|\mathbf{Z}\backslash Z_m\Big] = \begin{cases} \textbf{Random Variable} & \textbf{Probability} \\ Y_1^2 & F_\theta^+(Y_1) \\ \mathbb{E}_\theta\big[(Z_m^+)^2\big|Z_m^+ \in (Y_1, Y_2];\mathbf{Z}\backslash Z_m\big] & F_\theta^+(Y_2) - F_\theta^+(Y_1) \\ Y_2^2 & 1 - F_\theta^+(Y_2). \end{cases}$$

The conditional expectation $\mathbb{E}[(B^-(m))^2|\mathbf{Z}\backslash(Z_{m-1}, Z_m)]$ is similar. Applying the integration by parts method to the middle outcome in the above in distribution equality, we have

$$\mathbb{E}_\theta\big[(Z_m^+)^2\,\big|\,Z_m^+ \in (Y_1, Y_2];\mathbf{Z}\backslash Z_m\big] = Y_2^2 F_\theta^+(Y_2) + Y_1^2 F_\theta^+(Y_1) - \int_{Y_1}^{Y_2} 2zF_\theta^+(z)\,dz,$$

and after simplifications, the conditional second moment $\mathbb{E}_\theta[(B^+(m))^2|\mathbf{Z}\backslash Z_m]$ is akin to the first moment, which immediately follows (2.18):

$$\mathbb{E}_\theta\Big[\big(B^+(m)\big)^2\Big|\mathbf{Z}\backslash Z_m\Big] = Y_2^2 - \int_{Y_1}^{Y_2} 2zF_\theta^+(z)\,dz,$$

with the equivalent negative variant. To repeat and collate, the conditional mean, $\mathbb{E}_\theta[B^+(m)|\mathbf{Z}\backslash Z_m]$, is written as:

$$\mathbb{E}_\theta\big[B^+(m)\,\big|\mathbf{Z}\backslash Z_m\big] = Y_2 - \int_{Y_1}^{Y_2} F_\theta^+(z)\,dz. \tag{3.29}$$

The negative variant is analogous. After some algebra, incorporating these four results, the integral expression for the conditional second moment is:

$$\begin{aligned}&\mathbb{E}_\theta\Big[\big(B^+(m)-B^-(m)\big)^2\Big|\mathbf{Z}\backslash Z_m\Big]\\ &= \int_{Y_1}^{Y_2} 2(Y_2-z)\big(F_\theta^+(z)+F_\theta^-(z)\big)\,dz - 2\int_{Y_1}^{Y_2}\int_{Y_1}^{Y_2} F_\theta^+(z)F_\theta^-(w)\,dz\,dw.\\ &:= I_1 + I_2.\end{aligned}$$

We will compute I_1 and I_2 presently. To extract our limiting terms, we use a two term Taylor series for each integrand with Lagrange remainder, centred at $z = Y_1$, respectively $(z, w) = (Y_1, Y_1)$. For the first integral I_1 in the prior equation, our expansion yields, for given $\mathbf{Z}\backslash Z_m$:

$$\begin{aligned}I_1 = &\big(F_\theta^+(Y_1)+F_\theta^-(Y_1)\big)(Y_2-Y_1)^2 - \int_{Y_1}^{Y_2}\big(F_\theta^+(\xi_1)+F_\theta^-(\xi_1)\big)(z-Y_1)\,dz\\ &+\int_{Y_1}^{Y_2}\big(f_\theta^+(\xi_2)+f_\theta^-(\xi_2)\big)(Y_2-\xi_2)(z-Y_1)\,dz,\end{aligned}$$

with $\xi_1,\ \xi_2 \in (Y_1, Y_2) \subset B_{2r_{m-1}}(q_\alpha(\theta))$, defined in Lemma A.2. Since F_θ^+, F_θ^-, are increasing on $B_{2r_{m-1}}(q_\alpha(\theta))$, and noting that $Y_2-\xi_2 \le Y_2-Y_1$, this expression has upper bound

$$I_1 \le \frac{1}{2}\big(F_\theta^+(Y_1)+F_\theta^-(Y_1)\big)(Y_2-Y_1)^2 + \frac{1}{2}c_1(Y_2-Y_1)^3,$$

where $c_1 = \sup_{y\in B_{2r_{m-1}}(q_\alpha(\theta))}|f_\theta^+(y)+f_\theta^-(y)|$. The lower bound is attained from the same reasoning:

$$I_1 \ge \big(F_\theta^+(Y_1)+F_\theta^-(Y_1)\big)(Y_2-Y_1)^2 - \frac{1}{2}\big(F_\theta^+(Y_2)+F_\theta^-(Y_2)\big)(Y_2-Y_1)^2 - \frac{1}{2}c_1(Y_2-Y_1)^3.$$

Both upper and lower bounds are integrable from the uniform convergence of spacings, Lemma A.2, and the boundedness of the measure-valued derivative density functions, Assumption **(B6)**. We again use the same machinery as in Section 3.4.1 to attain the limiting result for $m \to \infty$. The order statistics Y_1, $Y_2 \overset{a.s.}{\to} q_\alpha(\theta)$, from Bahadur, [4], and $F_\theta^+(Y_l) + F_\theta^-(Y_l) \overset{a.s.}{\to} F_\theta^+(q_\alpha(\theta)) + F_\theta^-(\theta)$, $l = 1, 2$, by the Continuous Mapping Theorem. Together with the spacing limit from Pyke, [86], and the Slutsky and Sandwich Theorems, in distribution

$$m^2 I_1 \overset{d}{\to} \frac{F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))}{f_\theta(q_\alpha(\theta)} E^2, \tag{3.30}$$

in which E is an exponential mean one random variable. For the integral I_2 in (3.30), the two term Taylor expansion yields a conceptually simpler upper bound

$$I_2 \leq \left(F_\theta^+(Y_1)(Y_2 - Y_1) + c_2(Y_2 - Y_1)^2\right)\left(F_\theta^-(Y_1)(Y_2 - Y_1) + c_3(Y_2 - Y_1)\right)^2,$$

in which $c_2 = \sup_{y \in B_{2r_{m-1}}(q_\alpha(\theta))} f_\theta^+(y)$, and $c_3 = \sup_{y \in B_{2r_{m-1}}(q_\alpha(\theta))} f_\theta^-(y)$. For the lower bound, the infimums replace the supremums in the former constants. As the leading order of the spacing is two, $m^2 I_2$ is uniformly integrable. From the same arguments that preceded (3.30), in distribution:

$$m\left(F_\theta^+(Y_1)(Y_2 - Y_1) + c_2(Y_2 - Y_1)^2\right) \overset{d}{\to} \frac{F_\theta(q_\alpha(\theta))}{f_\theta(q_\alpha(\theta)} E,$$

as $m \to \infty$, with an analogous limit for the negative variant. The exponential limiting random variable is the same as in (3.30), originating from the same sequence. We note in particular that $m(Y_2 - Y_1)^2 \overset{a.s.}{\to} 0$ from part (ii) of Lemma A.2. A combination of the Continuous Mapping Theorem, with map $b(x, y) = xy$, and another application of Slutsky's Theorem provides the limit for I_2:

$$m^2 I_2 \overset{d}{\to} \frac{F_\theta^+(q_\alpha(\theta))F_\theta^-(q_\alpha(\theta))}{f_\theta^2(q_\alpha(\theta)} E^2 \tag{3.31}$$

as $m \to \infty$.

The sum of these two limits, Equations (3.30) and (3.31), noting that $\mathbb{E}[E^2] = 2$, provides us the limiting result for the second moment via the Dominated Convergence Theorem:

$$\begin{aligned}
&\lim_{m\to\infty} \mathbb{E}_\theta\Big[m^2\left(B^+(m) - B^-(m)\right)^2\Big] \\
&\quad = \mathbb{E}\Big[\lim_{m\to\infty} \mathbb{E}_\theta\Big[m^2\left(B^+(m) - B^-(m)\right)^2\Big|\mathbf{Z}\backslash Z_m\Big]\Big] \\
&\quad = \frac{2}{f_\theta^2(q_\alpha(\theta))}\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta) + F_\theta^+(q_\alpha(\theta))F_\theta^-(q_\alpha(\theta))\right).
\end{aligned}$$

Together with Equation (3.27), the asymptotic variance defined in Equation (3.26), after a minor rearrangement, yields

$$\lim_{m\to\infty} \mathrm{Var}_\theta(D_2^{\mathrm{MVD}}(m,k)) = \frac{c_\theta^2}{k f_\theta(q_\alpha(\theta))}\Big(2\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right) - \left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right)^2\Big), \tag{3.32}$$

which is the expression for part (i) of this lemma.

Part (ii): Supposing Assumption **(B3(c))**, the possible dependence between a measure-valued derivative random variable pair (Z^+, Z^-) includes two more terms into the conditional second moment, Equation (3.28). In addition, we also condition the random variables denoted as conditional expectations $\mathbb{E}_\theta[B^\pm(m)|\mathbf{Z}\backslash Z_m]$ w.r.t. the sequence $\mathbf{Z}$ to draw out the effect of the dependence present within at least one of the random variable pairs (Z_m^+, Z_m), (Z_m^-, Z_m):

$$\begin{aligned}\mathbb{E}_\theta\Big[\left(B^+(m) - B^-(m)\right)^2\Big|\mathbf{Z}\backslash Z_m\Big] = \mathbb{E}_\theta\Big[&\left(B^+_{\mathrm{IND}}(m) - B^-_{\mathrm{IND}}(m)\right)^2\Big|\mathbf{Z}\backslash Z_m\Big] \\ &+ 2\left(\mathbb{E}\left[\mathbb{E}_\theta\left[B^+(m)B^-(m)\middle|\mathbf{Z}\right]\middle|\mathbf{Z}\backslash Z_m\right]\right. \\ &\left.- \mathbb{E}_\theta\left[B^+_{\mathrm{IND}}(m)B^-_{\mathrm{IND}}(m)\middle|\mathbf{Z}\backslash Z_m\right]\right).\end{aligned}$$

This is the only conditional expectation in this derivation that is different from part (i). The subscript denotes that the calculation is evaluated assuming mutual independence. Specifically, the asymptotic variance is computed according to Equation (3.26) and we utilize the limiting expectation, Equation (3.27). The limiting expression, as $m \to \infty$, for the asymptotic variance contingent on Assumption **(B3(c))**, from (3.32), is then

$$\begin{aligned}&\lim_{m\to\infty} \mathrm{Var}_\theta\left(D_2^{\mathrm{MVD}}(m,k)\right) \\ &= \frac{c_\theta^2}{k}\bigg(\frac{1}{f_\theta(q_\alpha(\theta))}\Big(2\left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right) - \left(F_\theta^+(q_\alpha(\theta)) + F_\theta^-(q_\alpha(\theta))\right)^2\Big) \\ &\quad + \lim_{m\to\infty} 2\mathbb{E}\Big[\mathbb{E}\left[\mathbb{E}_\theta\left[m^2 B^+(m)B^-(m)\middle|\mathbf{Z}\right]\middle|\mathbf{Z}\backslash Z_m\right] - \mathbb{E}\left[m^2 B^+_{\mathrm{IND}}(m)B^-_{\mathrm{IND}}(m)\middle|\mathbf{Z}\backslash Z_m\right]\Big]\bigg).\end{aligned} \tag{3.33}$$

The possible dependence between Z^+ and Z^- prevents any early simplification of the conditional expectation $\mathbb{E}_\theta[B^+(m)B^-(m)|\mathbf{Z}]$ enjoyed for the proof of part (i). As discussed at the end of Section 3.2.4, the nature of these outcomes occur individually, when substituting Z_m for Z_m^+ and respectively Z_m^- to obtain $B^+(m) = Z^+_{\lceil \alpha m\rceil:m}$, $B^-(m) = Z^-_{\lceil \alpha m\rceil:m}$. In terms of the sample space where these nine outcomes appear, the with the almost sure equivalent outcomes for the random variable $B^+(m)B^-(m)$ given $\mathbf{Z}$ is written below:

$$B^+(m)B^-(m)\,\big|\,\mathbf{Z} = \begin{cases} \textbf{Random Variable} & \textbf{Event} \\ Y_1^2 & \{\omega : Z_m^+ \leq Y_1, Z_m^- \leq Y_1\} \\ Y_2^2 & \{\omega : Z_m^+ > Y_2, Z_m^- > Y_2\} \\ Y_1 Y_2 & \{\omega : Z_m^+ \leq Y_1, Z_m^- > Y_2\} \\ & \quad \cup \{\omega : Z_m^+ > Y_2, Z_m^- \leq Y_1\} \\ Y_1 Z_m^+ & \{\omega : Y_1 < Z_m^+ \leq Y_2, Z_m^- \leq Y_1\} \\ Y_2 Z_m^+ & \{\omega : Y_1 < Z_m^+ \leq Y_2, Z_m^- > Y_2\} \\ Y_1 Z_m^- & \{\omega : Z_m^+ \leq Y_1, Y_1 < Z_m^- \leq Y_2\} \\ Y_2 Z_m^- & \{\omega : Z_m^+ > Y_2, Y_1 < Z_m^- \leq Y_2\} \\ Z_m^+ Z_m^- & \{\omega : Y_1 < Z_m^+ \leq Y_2, Y_1 < Z_m^- \leq Y_2\}. \end{cases}$$

The diagram depicting these outcomes, Figure 3.1, was given in Section 3.2.4. We remind that the order statistics $Y_1 = Z_{\lceil \alpha m \rceil -1:m-1}$, $Y_2 = Z_{\lceil \alpha m \rceil :m}$ are formed from the sequence $\mathbf{Z} \backslash Z_m$.

To arrive at this conditional expectation, we begin by attaining expressions for each of the components before combining. These expectations are first conditioned on $\mathbf{Z}$, before we make use of the tower property to remove the conditioning w.r.t. $\mathbf{Z}$. In this section, $F_\theta^\pm$ is introduced to be the distribution function of the random variable pair (Z^+, Z^-) and $G_\theta^\pm$ for the trivariate distribution (Z^+, Z^-, Z). In addition we remind that, for instance, $F_\theta^+(\cdot|Z_m)$ is the regular conditional distribution function of Z_m^+ conditioned on Z_m. For all of the conditional random variables presented, a unique regular conditional distribution does exist from [62], p. 107, ensuring that these conditional distributions are also Stieltjes measures. Lastly, we require continuity in the neighbourhood of the value of the quantile for the bivariate distribution $F_\theta^\pm$. This is proven in Postscript I in Section 3.4.1, given the neighbourhood continuity of the associate marginal distributions.

From the nine equivalent outcomes of $B^+(m)B^-(m)$, we start at the topmost and increase in complexity.

$$\mathbb{E}_\theta\big[1\{Z^+ \leq Y_1\}\,1\{Z^- \leq Y_1\}\big|\mathbf{Z}\big] = \int_{(-\infty, Y_1] \times (-\infty, Y_1]} dF_\theta^\pm(u, v|Z_m)$$

and

$$\begin{aligned}\mathbb{E}_\theta\big[1\{Z^+ \le Y_1\}\,1\{Z^- \le Y_1\}\big|\mathbf{Z}\backslash Z_m\big] &= \mathbb{E}\big[\mathbb{E}_\theta\big[1\{Z^+ \le Y_1\}\,1\{Z^- \le Y_1\}\big|\mathbf{Z}\big]\big|\mathbf{Z}\backslash Z_m\big]\\ &= \int_{\mathbb{R}}\int_{(-\infty,Y_1]\times(-\infty,Y_1]} dF_\theta^{\pm}(u,v|w)\,dF_\theta(w)\\ &= \int_{(-\infty,Y_1]\times(-\infty,Y_1]}\int_{\mathbb{R}} dG_\theta^{\pm}(u,v,w).\end{aligned}$$

Then via the Disintegration Theorem, this is equal to

$$= \int_{(-\infty,Y_1]\times(-\infty,Y_1]} dF_\theta^{\pm}(u,v) \;=\; F_\theta^{\pm}(Y_1,Y_1) \tag{3.34}$$

In addition,

$$\mathbb{E}_\theta\big[1\{Z^+ > Y_2\}\,1\{Z^- > Y_2\}\big|\mathbf{Z}\big] = \int_{(Y_2,\infty)\times(Y_2,\infty)} dF_\theta^{\pm}(u,v|Z_m), \tag{3.35}$$

and by repeating the argument

$$\mathbb{E}_\theta\big[1\{Z^+ > Y_2\}\,1\{Z^- > Y_2\}\big|\mathbf{Z}\backslash Z_m\big] = \int_{(Y_2,\infty)\times(Y_2,\infty)} dF_\theta^{\pm}(u,v).$$

Since $F_\theta^{\pm}$ is continuous on $B^2_{2r_{m-1}}(q_\alpha(\theta))$, this integral can be written in the standard form

$$\begin{aligned}&= \int_{\mathbb{R}^2} dF_\theta^{\pm}(u,v) - \int_{(-\infty,Y_2]\times\mathbb{R}} dF_\theta^{\pm}(u,v) - \int_{\mathbb{R}\times(-\infty,Y_2]} dF_\theta^{\pm}(u,v)\\ &\quad + \int_{(-\infty,Y_2]\times(-\infty,Y_2]} dF_\theta^{\pm}(u,v)\\ &= 1 - F_\theta^{+}(Y_2) - F_\theta^{-}(Y_2) + F_\theta^{\pm}(Y_2,Y_2).\end{aligned} \tag{3.36}$$

Similarly,

$$\begin{aligned}&\mathbb{E}_\theta\big[1\{Z^+ \le Y_1\}\,1\{Z^- > Y_2\} + 1\{Z^+ > Y_2\}\,1\{Z^- \le Y_1\}\big|\mathbf{Z}\backslash Z_m\big]\\ &\quad = F_\theta^{+}(Y_1) - F_\theta^{\pm}(Y_1,Y_2) + F_\theta^{-}(Y_1) - F_\theta^{\pm}(Y_2,Y_1).\end{aligned} \tag{3.37}$$

For the conditional moments, the final integrands, via the Lebesgue-Stieltjes integration by parts formula, [96], are depicted via its left limit, for instance $F^+(z^-)$. We need the assumed continuity on $B_{2r_{m-1}}(q_\alpha(\theta))$ for the marginal distributions and proven continuity in $B^2_{2r_{m-1}}(q_\alpha(\theta))$, via Postscript I in Section 3.4.1, to remove the minus sign from the arguments of the bivariate distributions. From this point onwards, we proceed from the position where the condition w.r.t. Z_m has been removed. The method is identical to (3.34) and (3.36):

$$
\begin{aligned}
&\mathbb{E}_\theta\big[Z_m^+ \, 1\{Y_1 < Z_m^+ \le Y_2\} \, 1\{Z_m^- \le Y_1\} \big| \mathbf{Z} \backslash Z_m\big] \\
&\quad = \int_{(Y_1,Y_2]\times(-\infty,Y_1]} u \, dF_\theta^\pm(u,v) \\
&\quad = \int_{(Y_1,Y_2]\times(-\infty,Y_1]} u \, dF_\theta^\pm(u|v) \, dF_\theta^-(v),
\end{aligned}
$$

and via the Distintegration Theorem,

$$
\begin{aligned}
&= \int_{(-\infty,Y_1]} \left(Y_2 F_\theta^+(Y_2|v) - Y_1 F_\theta^+(Y_1|v) - \int_{(Y_1,Y_2]} dF_\theta^+(u^-|v) \right) dF_\theta^-(v) \\
&= Y_2 \int_{(-\infty,Y_2]\times(-\infty,Y_1]} dF_\theta^+(u|v) \, dF_\theta^-(v) - Y_1 \int_{(-\infty,Y_1]\times(-\infty,Y_1]} dF_\theta^+(u|v) \, dF_\theta^-(v) \\
&\quad - \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]\times(-\infty,Y_1]} dF_\theta^\pm(s,v) \, du \\
&= Y_2 F_\theta^\pm(Y_2,Y_1) - Y_1 F_\theta^\pm(Y_1,Y_1) - \int_{Y_1}^{Y_2} F_\theta^\pm(u,Y_1) \, du, \qquad (3.38)
\end{aligned}
$$

as $F_\theta^\pm$ is continuous on $B^2_{2r_{m-1}}(q_\alpha(\theta))$. Correspondingly,

$$
\begin{aligned}
&\mathbb{E}_\theta\big[Z_m^- \, 1\{Z_m^+ \le Y_1\} \, 1\{Y_1 < Z_m^- \le Y_2\} \big| \mathbf{Z} \backslash Z_m\big] \\
&\quad = Y_2 F_\theta^\pm(Y_1,Y_2) - Y_1 F_\theta^\pm(Y_1,Y_1) - \int_{Y_1}^{Y_2} F_\theta^\pm(Y_1,v) \, dv. \qquad (3.39)
\end{aligned}
$$

For the seventh and eighth outcomes, there are only minor differences in the derivation:

$$
\begin{aligned}
&\mathbb{E}_\theta\big[Z_m^+ \, 1\{Y_1 < Z_m^+ \le Y_2\} \, 1\{Z_m^- > Y_2\} \big| \mathbf{Z} \backslash Z_m\big] \\
&\quad = \int_{(Y_1,Y_2]\times(Y_2,\infty)} u \, dF_\theta^+(u,v) \\
&\quad = \int_{(Y_1,Y_2]\times(Y_2,\infty)} u \, dF_\theta^+(u|v) \, dF_\theta^-(v),
\end{aligned}
$$

and via the Disintegration Theorem,

$$
= \int_{(Y_2,\infty)} \left(Y_2 F_\theta^+(Y_2|v) - Y_1 F_\theta^+(Y_1|v) - \int_{(Y_1,Y_2]} dF_\theta^\pm(u|v) \right) dF_\theta^-(v)
$$

$$
\begin{aligned}
&= Y_2 \int_{(-\infty,Y_2]\times(Y_2,\infty)} dF_\theta^{\pm}(u,v) - Y_1 \int_{(-\infty,Y_1]\times(Y_2,\infty)} dF_\theta^{\pm}(u,v) - \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]\times(Y_2,\infty)} dF_\theta^{\pm}(s,v)\,du \\
&= Y_2 \left(F_\theta^{+}(Y_2) - F_\theta^{\pm}(Y_2,Y_2)\right) - Y_1 \left(F_\theta^{+}(Y_1) - F_\theta^{\pm}(Y_1,Y_2)\right) - \int_{Y_1}^{Y_2} F_\theta^{+}(u)\,du \\
&\quad + \int_{Y_1}^{Y_2} F_\theta^{\pm}(u,Y_2)\,du, \qquad (3.40)
\end{aligned}
$$

since $F_\theta^{\pm}$ is continuous on $B^2_{2r_{m-1}}(q_\alpha(\theta))$. Likewise,

$$
\begin{aligned}
&\mathbb{E}_\theta\left[Z_m^- 1\{Z_m^+ > Y_2\}\, 1\{Y_1 < Z_m^- \le Y_2\} \middle| \mathbf{Z}\backslash Z_m\right] \\
&\quad = Y_2 \left(F_\theta^{-}(Y_2) - F_\theta^{\pm}(Y_2,Y_2)\right) - Y_1 \left(F_\theta^{-}(Y_1) - F_\theta^{\pm}(Y_2,Y_1)\right) - \int_{Y_1}^{Y_2} F_\theta^{-}(v)\,dv \\
&\quad + \int_{Y_1}^{Y_2} F_\theta^{\pm}(Y_2,v)\,dv. \qquad (3.41)
\end{aligned}
$$

For the final conditional expectation, integration by parts formula is used twice. The method is the same as earlier computations though slightly more intricate. By the Disintegration Theorem:

$$
\begin{aligned}
&\mathbb{E}_\theta\left[Z_m^+ Z_m^- 1\{Y_1 < Z_m^+ \le Y_2\}\, 1\{Y_1 < Z_m^- \le Y_2\} \middle| \mathbf{Z}\backslash Z_m\right] \\
&= \int_{(Y_1,Y_2]\times(Y_1,Y_2]} uv\, dF_\theta^{\pm}(u,v) \\
&= \int_{(Y_1,Y_2]} v \left(\int_{(Y_1,Y_2]} u\, dF_\theta^{\pm}(u|v) \right) dF_\theta^{-}(v) \\
&= \int_{(Y_1,Y_2]} v \left(Y_2 F_\theta^{\pm}(Y_2|v) - Y_1 F_\theta^{\pm}(Y_1|v) - \int_{(Y_1,Y_2]} dF_\theta^{\pm}(u^-|v) \right) dF_\theta^{-}(v) \\
&= Y_2 \int_{(-\infty,Y_2]\times(Y_1,Y_2]} v\, dF_\theta^{\pm}(u,v) - Y_1 \int_{(-\infty,Y_1]\times(Y_1,Y_2]} v\, dF_\theta^{\pm}(u,v) - \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]\times(Y_1,Y_2]} v\, dF_\theta^{\pm}(s,v)\,du,
\end{aligned}
$$

due to Fubini's Theorem. Since the bivariate measure can be disintegrated in the reverse arrangement, we have:

$$
\begin{aligned}
&= Y_2 \int_{(-\infty,Y_2]} \left(Y_2 F_\theta^{-}(Y_2|u) - Y_1 F_\theta^{-}(Y_1|u) - \int_{(Y_1,Y_2]} F_\theta^{-}(v|u)\,dv \right) dF_\theta^{+}(u) \\
&- Y_1 \int_{(-\infty,Y_1]} \left(Y_2 F_\theta^{-}(Y_2|u) - Y_1 F_\theta^{-}(Y_1|u) - \int_{(Y_1,Y_2]} F_\theta^{-}(v|u)\,dv \right) dF_\theta^{+}(u) \\
&- \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]} \left(Y_2 F_\theta^{-}(Y_2|s) - Y_1 F_\theta^{-}(Y_1|s) - \int_{(Y_1,Y_2]} F_\theta^{-}(v|s)\,dv \right) dF_\theta^{+}(s)\,du
\end{aligned}
$$

$$
\begin{aligned}
&= Y_2^2 \int_{(-\infty,Y_2]\times(-\infty,Y_2]} dF_\theta^{\pm}(u,v) - Y_1 Y_2 \int_{(-\infty,Y_2]\times(-\infty,Y_1]} dF_\theta^{\pm}(u,v) - Y_2 \int_{(Y_1,Y_2]} \int_{(-\infty,Y_2]\times(-\infty,v^-]} dF_\theta^{\pm}(u,t)\,dv \\
&\quad - Y_1 Y_2 \int_{(-\infty,Y_1]\times(-\infty,Y_2]} dF_\theta^{\pm}(u,v) + Y_1^2 \int_{(-\infty,Y_1]\times(-\infty,Y_1]} dF_\theta^{\pm}(u,v) + Y_1 \int_{(Y_1,Y_2]} \int_{(-\infty,Y_1]\times(-\infty,v^-]} dF_\theta^{\pm}(u,t)\,dv \\
&\quad - Y_2 \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]\times(-\infty,Y_2]} dF_\theta^{\pm}(s,v)\,du + Y_1 \int_{(Y_1,Y_2]} \int_{(-\infty,u^-]\times(-\infty,Y_1]} dF_\theta^{\pm}(s,v)\,du \\
&\quad + \int_{(Y_1,Y_2]\times(Y_1,Y_2]} \int_{(-\infty,u^-]\times(-\infty,v^-]} dF_\theta^{\pm}(s,t)\,du\,dv \\
&= Y_2^2 F_\theta^{\pm}(Y_2,Y_2) - Y_1 Y_2 F_\theta^{\pm}(Y_2,Y_1) - Y_2 \int_{Y_1}^{Y_2} F_\theta^{\pm}(Y_2,v)\,dv \\
&\quad - Y_1 Y_2 F_\theta^{\pm}(Y_1,Y_2) + Y_1^2 F_\theta^{\pm}(Y_1,Y_1) + Y_1 \int_{Y_1}^{Y_2} F_\theta^{\pm}(Y_1,v)\,dv \\
&\quad - Y_2 \int_{Y_1}^{Y_2} F_\theta^{\pm}(u,Y_2)\,du + Y_1 \int_{Y_1}^{Y_2} F_\theta^{\pm}(u,Y_1)\,du + \int_{Y_1}^{Y_2}\int_{Y_1}^{Y_2} F_\theta^{\pm}(u,v)\,du\,dv, \qquad (3.42)
\end{aligned}
$$

by continuity on $F_\theta^{\pm}$ in $B^2_{2r_{m-1}}(q_\alpha(\theta))$. When added, nearly all of the terms from Equations (3.34) - (3.42) cancel. The resulting expectation, $\mathbb{E}_\theta[B^+(m)B^-(m)|\mathbf{Z}\backslash Z_m]$, is then

$$
\begin{aligned}
&\mathbb{E}_\theta\big[B^+(m)B^-(m)\big|\mathbf{Z}\backslash Z_m\big] \\
&\quad = Y_1^2 \mathbb{E}_\theta\big[1\{Z^+ \le Y_1\}\,1\{Z^- < Y_1\}\big|\mathbf{Z}\backslash Z_m\big] + Y_2^2 \mathbb{E}_\theta\big[1\{Z^+ > Y_2\}\,1\{Z^- > Y_2\}\big|\mathbf{Z}\backslash Z_m\big] \\
&\qquad + Y_1 Y_2 \mathbb{E}_\theta\big[1\{Z^+ \le Y_1\}\,1\{Z^- > Y_2\} + 1\{Z^+ > Y_2\}\,1\{Z^- \le Y_1\}\big|\mathbf{Z}\backslash Z_m\big] \\
&\qquad + Y_1 \big(\mathbb{E}_\theta\big[Z_m^+\,1\{Y_1 < Z_m^+ \le Y_2\}\,1\{Z_m^- \le Y_1\}\big|\mathbf{Z}\backslash Z_m\big] \\
&\qquad\qquad + \mathbb{E}_\theta\big[Z_m^-\,1\{Z_m^+ > Y_2\}\,1\{Y_1 < Z_m^- \le Y_2\}\big|\mathbf{Z}\backslash Z_m\big]\big) \\
&\qquad + Y_2 \big(\mathbb{E}_\theta\big[Z_m^+\,1\{Y_1 < Z_m^+ \le Y_2\}\,1\{Z_m^- > Y_2\}\big|\mathbf{Z}\backslash Z_m\big] \\
&\qquad\qquad + \mathbb{E}_\theta\big[Z_m^-\,1\{Z_m^+ > Y_2\}\,1\{Y_1 < Z_m^- \le Y_2\}\big|\mathbf{Z}\backslash Z_m\big]\big) \\
&\qquad\quad + \mathbb{E}_\theta\big[Z_m^+ Z_m^-\,1\{Y_1 < Z_m^+ \le Y_2\}\,1\{Y_1 < Z_m^- \le Y_2\}\big|\mathbf{Z}\backslash Z_m\big],
\end{aligned}
$$

and after the cancellations, is equal to

$$
= Y_2^2 - Y_2 \int_{Y_1}^{Y_2} F_\theta^+(u)\,du - Y_2 \int_{Y_1}^{Y_2} F_\theta^-(v)\,dv + \int_{Y_1}^{Y_2}\int_{Y_1}^{Y_2} F_\theta^{\pm}(u,v)\,du\,dv.
$$

From Equation (3.29), the conditional expectation contingent on independent elements within (Z^+, Z^-) yields

$$
\begin{aligned}
\mathbb{E}_\theta\big[B^+_{\text{IND}}(m)B^-_{\text{IND}}(m)\,\big|\mathbf{Z}\backslash Z_m\big] &= Y_2^2 - Y_2 \int_{Y_1}^{Y_2} \big(F_\theta^+(u) + F_\theta^-(u)\big)\,du \\
&\quad + \int_{Y_1}^{Y_2}\int_{Y_1}^{Y_2} F_\theta^+(u) F_\theta^-(v)\,du\,dv,
\end{aligned}
$$

and so the difference between the previous two expressions is the difference in the double integrals

$$\begin{aligned}&\mathbb{E}_\theta\big[m^2B^+(m)B^-(m)\,\big|\,\mathbf{Z}\backslash Z_m\big]-\mathbb{E}_\theta\big[m^2B^+_{\text{IND}}(m)B^-_{\text{IND}}(m)\,\big|\,\mathbf{Z}\backslash Z_m\big]\\&=\int_{Y_1}^{Y_2}\int_{Y_1}^{Y_2}m^2\left(F^{\pm}_\theta(u,v)-F^+_\theta(u)F^-_\theta(v)\right)du\,dv.\end{aligned}$$

And this expression has an upper bound of

$$\leq m^2\left(F^{\pm}_\theta(Y_2,Y_2)-F^+_\theta(Y_1)F^-_\theta(Y_1)\right)(Y_2-Y_1)^2$$

due to the distribution function being non-negative and non-decreasing. The lower bound occurs by a vice versa replacement of Y_2 and Y_1 within the distribution functions. This expression is uniformly integrable via part (i) of Lemma A.2 together with the fact that the difference of distribution functions is bounded absolutely by one. As $m\to\infty$, from [4], $Y_1,Y_2\overset{a.s.}{\to}q_\alpha(\theta)$, and, by the Continuous Mapping Theorem, an analogous almost sure result is attained for each distribution function. Jointly with another application of the Continuous Mapping Theorem, with map $a(x)=x^2$, the limiting spacing result yields $m^2(Y_2-Y_1)^2\overset{d}{\to}E^2/f^2_\theta(q_\alpha(\theta))$. Then from Slutsky's Theorem, combining all the limits together,

$$\begin{aligned}&\mathbb{E}_\theta\big[m^2B^+(m)B^-(m)\big|\mathbf{Z}\backslash Z_m\big]-\mathbb{E}_\theta\big[m^2B^+_{\text{IND}}(m)B^-_{\text{IND}}(m)\big|\mathbf{Z}\backslash Z_m\big]\overset{d}{\to}\\&\quad\frac{F^{\pm}_\theta(q_\alpha(\theta),q_\alpha(\theta))-F^+_\theta(q_\alpha(\theta))F^-_\theta(q_\alpha(\theta))}{f^2_\theta(q_\alpha(\theta))}E^2,\end{aligned}$$

as $m\to\infty$, and via the Dominated Convergence Theorem, with $\mathbb{E}[E^2]=2$

$$\begin{aligned}&\lim_{m\to\infty}\mathbb{E}_\theta\big[m^2B^+(m)B^-(m)\big]-\mathbb{E}_\theta\big[m^2B^+_{\text{IND}}(m)B^-_{\text{IND}}(m)\big]\\&=\mathbb{E}\Big[\lim_{m\to\infty}\mathbb{E}_\theta\big[m^2B^+(m)B^-(m)\big|\mathbf{Z}\backslash Z_m\big]-\mathbb{E}_\theta\big[m^2B^+_{\text{IND}}(m)B^-_{\text{IND}}(m)\big|\mathbf{Z}\backslash Z_m\big]\Big]\\&=\frac{2\Big(F^{\pm}_\theta(q_\alpha(\theta),q_\alpha(\theta))-F^+_\theta(q_\alpha(\theta))F^-_\theta(q_\alpha(\theta))\Big)}{f^2_\theta(q_\alpha(\theta))}.\end{aligned}$$

The substitution of this result into Equation (3.33) then provides us with the asymptotic variance in part (ii).

3.4.3 Importance Sampling

For either derivative estimator, $l=1,2$, the expectation, focusing on the general MVD ranked-data estimator, from Theorem 2.2 in Chapter 2 reduces to the expression

$$\mathbb{E}_\theta[D_l^{\mathrm{MVD}}(m,k)] = \mathbb{E}\Big[mc_\theta\left(Z^+_{\lceil\alpha m\rceil:m} - Z^-_{\lceil\alpha m\rceil:m}\right)\Big].$$

We utilize once more, see Section 3.2.2, that $Y_1 = Z_{\lceil\alpha m\rceil-1:m-1}$ and $Y_2 = Z_{\lceil\alpha m\rceil:m-1}$ are composed from the $m-1$ random variables in the sequence $\mathbf{Z}\backslash Z_m$. For the importance sampling version of this estimator, from Section 3.2.6 preceding Equation (3.13), the inclusion of the constructed importance sampling random variables $\check{Z}^+_m$, $\check{Z}^-_m$, respectively, into $\mathbf{Z}\backslash Z_m$ to form the associate order statistics $\check{Z}^+_{\lceil\alpha m\rceil:m}$, $\check{Z}^-_{\lceil\alpha m\rceil:m}$ are the random variables themselves:

$$\begin{aligned} &\check{Z}^+_{\lceil\alpha m\rceil:m} | \mathbf{Z}\backslash Z_m = \check{Z}^+_m, \quad \text{and} \\ &\check{Z}^-_{\lceil\alpha m\rceil:m} | \mathbf{Z}\backslash Z_m = \check{Z}^-_m. \end{aligned}$$

The associated expectations of these random variables are attained via integration by parts. For the positive term, $\mathbb{E}_\theta[Z^+_{\lceil\alpha m\rceil:m}|\mathbf{Z}\backslash Z_m]$, this is equal to:

$$\begin{aligned} \mathbb{E}_\theta\Big[\check{Z}^+_{\lceil\alpha m\rceil:m}\Big|\mathbf{Z}\backslash Z_m\Big] &= \mathbb{E}_\theta\big[\check{Z}^+_m\big] \\ &:= \mathbb{E}_\theta\big[Z^+_m \,\big|\, Z^+_m \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] \\ &= \int_{Y_1}^{Y_2} z \frac{f^+_\theta(z)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)}\, dz \\ &= \frac{Y_2 F^+_\theta(Y_2) - Y_1 F^+_\theta(Y_1)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)} - \int_{Y_1}^{Y_2} \frac{F^+_\theta(z)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)}\, dz \\ &= Y_2 + \frac{F^+_\theta(Y_1)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)}(Y_2 - Y_1) - \int_{Y_1}^{Y_2} \frac{F^+_\theta(z)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)}\, dz \\ &= Y_2 - \int_{Y_1}^{Y_2} \frac{F^+_\theta(z) - F^+_\theta(Y_1)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)}\, dz. \end{aligned}$$

An analogous result is attained for the conditional expectation w.r.t. the integrand Z^-_m. The following expression is the basis of our result:

$$\mathbb{E}_\theta\big[Z^+_m - Z^-_m \big| Z^+_m, Z^-_m \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] = -\int_{Y_1}^{Y_2} \frac{F^+_\theta(z) - F^+_\theta(Y_1)}{F^+_\theta(Y_2) - F^+_\theta(Y_1)} - \frac{F^-_\theta(z) - F^-_\theta(Y_1)}{F^-_\theta(Y_2) - F^-_\theta(Y_1)}\, dz.$$

For the above integral, we require a three term Taylor expansion centred around $z = Y_1$. The first term in either expansion immediately equals zero. We implement a Lagrangian remainder with the associated random variables $\xi_1, \xi_2 \in$

$(Y_1, Y_2] \subset B_{2r_{m-1}}(q_\alpha(\theta))$. This implementation yields:

$$\begin{aligned}
&\mathbb{E}_\theta\big[Z_m^+ - Z_m^- \big| Z_m^+, Z_m^- \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] \\
&\quad = -\int_{Y_1}^{Y_2} \left(\frac{f_\theta^+(Y_1)}{F_\theta^+(Y_2) - F_\theta^+(Y_1)} - \frac{f_\theta^-(Y_1)}{F_\theta^-(Y_2) - F_\theta^-(Y_1)} \right) (z - Y_1)\, dz \\
&\qquad - \int_{Y_1}^{Y_2} \left(\frac{\frac{\partial}{\partial z} f_\theta^+(\xi_1)}{F_\theta^+(Y_2) - F_\theta^+(Y_1)} - \frac{\frac{\partial}{\partial z} f_\theta^-(\xi_2)}{F_\theta^-(Y_2) - F_\theta^-(Y_1)} \right) (x - Y_1)^2\, dz.
\end{aligned}$$

Denoting $c_1 = \sup_{y \in B_{2r_{m-1}}(q_\alpha(\theta))} |\partial_x f_\theta^+(y)|$, and $c_2 = \sup_{y \in B_{2r_{m-1}}(q_\alpha(\theta))} |\partial_x f_\theta^-(y)|$, this last expression has an upper bound of

$$\begin{aligned}
&\mathbb{E}_\theta\big[m c_\theta \left(Z_m^+ - Z_m^-\right) \big| Z_m^+, Z_m^- \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] \\
&\quad \le -\frac{1}{2} m c_\theta \left(\frac{f_\theta^+(Y_1)}{F_\theta^+(Y_2) - F_\theta^+(Y_1)} - \frac{f_\theta^-(Y_1)}{F_\theta^-(Y_2) - F_\theta^-(Y_1)} \right) (Y_2 - Y_1)^2 \\
&\qquad + m c_\theta \left(\frac{c_1}{F_\theta^+(Y_2) - F_\theta^+(Y_1)} + \frac{c_2}{F_\theta^-(Y_2) - F_\theta^-(Y_1)} \right) (Y_2 - Y_1)^3.
\end{aligned}$$

Correspondingly, the plus sign in front of the second term is replaced with a negative sign for the lower bound. From the Mean-Value Theorem, $F_\theta^+(Y_2) - F_\theta^-(Y_1) = f_\theta^+(\xi_3)\,(Y_2 - Y_1)$, with $\xi_3 \in (Y_1, Y_2] \subset B_{2r_{m-1}}(q_\alpha(\theta))$, and a similar result is attained for the negative variant where ξ_4 is an element in the same neighbourhood. The upper bound is now the clearer expression:

$$\begin{aligned}
&\mathbb{E}_\theta\big[m c_\theta \left(Z_m^+ - Z_m^-\right) \big| Z_m^+, Z_m^- \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] \\
&\quad \le -\frac{1}{2} m c_\theta \left(\frac{f_\theta^+(Y_1)}{f_\theta^+(\xi_3)} - \frac{f_\theta^-(Y_1)}{f_\theta^-(\xi_4)} \right) (Y_2 - Y_1) + m c_\theta \left(\frac{c_1}{f_\theta^+(\xi_3)} + \frac{c_2}{f_\theta^-(\xi_4)} \right) (Y_2 - Y_1)^2.
\end{aligned}$$

All of the forthcoming results are given as $m \to \infty$. From Lemma A.2, part (ii), and Slutsky's Theorem, [45], the second term in the above expression is zero almost surely. For the first term, $Y_1, Y_2 \overset{a.s.}{\to} q_\alpha(\theta)$ from Bahadur, [4], and since $\xi_l \in (Y_1, Y_2]$, $l = 3, 4$, each of these random variables also almost surely equals $q_\alpha(\theta)$. Since f_θ^+, f_θ^- are continuous in $B_{2r_{m-1}}(q_\alpha(\theta))$, $f_\theta^+(Y_1)/f_\theta^+(\xi_3) \overset{a.s.}{\to} 1$, and $f_\theta^-(Y_1)/f_\theta^-(\xi_4) \overset{a.s.}{\to} 1$. Hence, from Slutsky's Theorem, [45], the difference in densities expression almost surely equals zero. And given the distributional spacing limit of Pyke, [86], another application of Slutsky's Theorem yields a distributional upper bound of zero as $m \to \infty$. The lower bound obtains the same distributional limit following the same arguments. Therefore, from the Sandwich Theorem:

$$\mathbb{E}_\theta\big[m c_\theta \left(Z_m^+ - Z_m^-\right) \big| Z_m^+, Z_m^- \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big] \overset{d}{\to} 0,$$

as m tends to infinity. The upper and lower bound of this conditional expectation are uniformly integrable due to a combination of Lemma A.2, and Assumptions **(B6)** - **(B8)**. Then via the Dominated Convergence Theorem, the asymptotic expectation of the importance sampling estimator is

$$\begin{aligned}\lim_{m\to\infty} \mathbb{E}_\theta\big[\check{D}_l^{\mathrm{MVD}}(m,k)\big] &= \mathbb{E}\Big[\lim_{m\to\infty} \mathbb{E}_\theta\Big[mc_\theta\Big(\check{Z}^+_{\lceil\alpha m\rceil:m} - \check{Z}^-_{\lceil\alpha m\rceil:m}\Big)\Big|\mathbf{Z}\backslash Z_m\Big]\Big]\\ &= \mathbb{E}\Big[\lim_{m\to\infty} \mathbb{E}_\theta\big[mc_\theta\big(Z_m^+ - Z_m^-\big)\,\big|\, Z_m^+, Z_m^- \in (Y_1, Y_2]; \mathbf{Z}\backslash Z_m\big]\Big]\\ &= 0,\end{aligned}$$

and by the Strong Law of Large Numbers $\check{D}_l^{\mathrm{MVD}}(m,k) \overset{a.s.}{\to} 0$ as $(m,k)\to(\infty,\infty)$.