

ESTIMATING NUMBER DENSITY N_V – A COMPARISON OF AN IMPROVED SALTYKOV ESTIMATOR AND THE DISECTOR METHOD

ASHOT DAVTIAN¹, UTE HAHN², JOACHIM OHSER³, DIETRICH STOYAN¹

¹Institut für Stochastik, TU Bergakademie Freiberg, 09596 Freiberg, Germany; ²Department of Mathematics and Statistics, The University of Western Australia, Nedlands WA 6907, Australia; ³Institut für Techno- und Wirtschaftsmathematik, 67663 Kaiserslautern, Germany
(Accepted March 17, 2000)

ABSTRACT

Two methods for the estimation of number per unit volume N_V of spherical particles are discussed: the (physical) disector (Sterio, 1984) and Saltykov's estimator (Saltykov, 1950; Fullman, 1953). A modification of Saltykov's estimator is proposed which reduces the variance. Formulae for bias and variance are given for both disector and improved Saltykov estimator for the case of randomly positioned particles. They enable the comparison of the two estimators with respect to their precision in terms of mean squared error.

Keywords: disector, number density, planar sections, precision, Saltykov estimator

INTRODUCTION

Estimating the mean number of particles per unit volume N_V is a fundamental stereological problem. There are two main approaches, the choice of which is mostly driven by practical considerations.

In cases where the structure of interest can be investigated as a three-dimensional structure, as e.g. in confocal microscopy, the (optical) disector is the most convenient method, and it gives unbiased estimates of N_V (Sterio, 1984; Howard *et al.*, 1985). Therefore it is very well established in biological sciences. However if the matter is opaque, the physical disector has to be used which is based on parallel sections a known distance apart. Then it is necessary to identify particles that are hit by both sections, which requires perfect registration of the relative position of the two images. Moreover a bias may be introduced because particles that are smaller than the distance between the two section planes cannot be detected. Both optical and physical disector imply a lot of manual work of qualified assistants.

Often it is practically very difficult or even impossible to make parallel sections, as of hard and brittle material. This is a reason why classical stereological methods using single sections are traditionally preferred e.g. by materials scientists. These methods are based on assumptions about the shape of the particles. Perhaps the most popular estimator is Saltykov's method (Saltykov, 1950; Fullman, 1953). It assumes that the particles are spherical, an assumption which in engineering is very often quite realistic. Saltykov's estimator requires only to measure the diameters of the (circular) particle

profiles. In many cases this can be accomplished by automated image analysis. The estimator is often blamed for its theoretically infinite variance, which is caused by very small intersection circles. As shown in this paper, it can easily be improved by binning these small diameters. Then its variance is finite at the cost of a negligible bias.

Up to now, hardly anything has been known about the statistical properties of the physical disector and the improved Saltykov estimator. It is clear that the variance of the estimators largely depends on the degree of regularity of the investigated structure, however the relationship is very complex, and in practical applications it is difficult to investigate the spatial variability of particle arrangement. We give formulae for variance and bias of both methods based on the assumptions that the particles are spherical and completely randomly positioned. Accuracy of the estimators is computed for various distributions of sphere diameters. The results lead to practical recommendations for the observation window size which is necessary to obtain a given mean squared error.

ESTIMATORS OF N_V

A well-known classical planar-section estimator \hat{N}_V^S of N_V (Saltykov, 1950; Fullman, 1953) is based on the observation of section circles with centres in a planar window W of area $A(W)$. Their (random) number is n , their diameters are s_1, \dots, s_n . Then

$$\hat{N}_V^S = \frac{2}{\pi A(W)} \sum_{i=1}^n \frac{1}{s_i}. \quad (1)$$

This estimator is unbiased, given the section plane is randomly positioned or, in terms of the model-approach, the system of spheres is homogeneous, i.e. can be described by a stationary germ grain model (Stoyan *et al.*, 1995). Unfortunately, the variance of \hat{N}_V^S is infinite due to the occurrence of very small section circles. In practice microscopical image resolution sets a natural lower limit to the diameters that can be observed, i.e. there is a lower bound for the s_i . But still the variance of \hat{N}_V^S will be very large.

Thus it is natural to modify the estimator such that very small section circles are not measured but only counted. To this purpose a lower diameter ε is introduced such that all section circle diameters larger than ε are measured while all smaller ones are set equal $\varepsilon/2$. With the notation

$$s_i^*(\varepsilon) = \begin{cases} \varepsilon/2 & \text{if } s_i < \varepsilon, \\ s_i & \text{otherwise} \end{cases}$$

the modified estimator is

$$\hat{N}_V^{S*} = \frac{2}{\pi A} \sum_{i=1}^n \frac{1}{s_i^*(\varepsilon)}. \quad (2)$$

Clearly, \hat{N}_V^{S*} must have a bias which increases with increasing ε . As shown below, the bias is rather small and the variance of \hat{N}_V^{S*} is finite, decreasing with increasing ε .

The physical disector estimator \hat{N}_V^D (Sterio, 1984) is based on counts made on two parallel section planes of distance t . The number Q^- of particles which appear in a counting frame in the reference plane but do not hit the look up plane has to be determined. With t denoting the distance of the two planes and $A(W)$ the area of the counting frame,

$$\hat{N}_V^D = \frac{Q^-}{tA(W)}. \quad (3)$$

This estimator is unbiased only if there are not any particles with diameter smaller than t . Otherwise such small particles may be situated between the two planes and therefore are overlooked, which results in a negative bias. Furthermore, the disector requires the distance t to be precisely known and that it is possible to decide whether profiles in the two planes belong to the same particle or not. The same assumptions about homogeneity of the particle system or randomization of the section planes are made as for the estimators \hat{N}_V^S and \hat{N}_V^{S*} .

STATISTICAL PROPERTIES OF THE N_V ESTIMATORS

Under the homogeneity assumptions made in the previous section, the means and biases of \hat{N}_V^{S*} and \hat{N}_V^D can be calculated quite easily. In this section only results are presented, for sketches of their derivation see the Appendix.

Means (and biases) of the estimators depend on the distribution function D_V of the sphere diameters. In the Appendix it is shown that

$$E\hat{N}_V^{S*} = N_V \left(\frac{2}{\pi} \int_{\varepsilon}^{\infty} \arccos\left(\frac{\varepsilon}{d}\right) D_V(dd) + \frac{4}{\pi\varepsilon} \int_0^{\infty} (D_V(\sqrt{u^2 + \varepsilon^2}) - D_V(u)) du \right) \quad (4)$$

and

$$E\hat{N}_V^D = N_V \left(1 - \frac{1}{t} \int_0^t (t-d) D_V(dd) \right). \quad (5)$$

Using these formulae the *relative bias* $B' = (E\hat{N}_V - N_V)/N_V$ was calculated for the two estimators. Five diameter distributions were considered, all with mean one: constant diameters, uniform (on [0.7,1.3]) and triangular (on [0.1,1.9]) as examples of bounded distributions, as well as Rayleigh distribution and lognormal distribution (with variance 0.1). The corresponding density functions are depicted in Fig. 1.

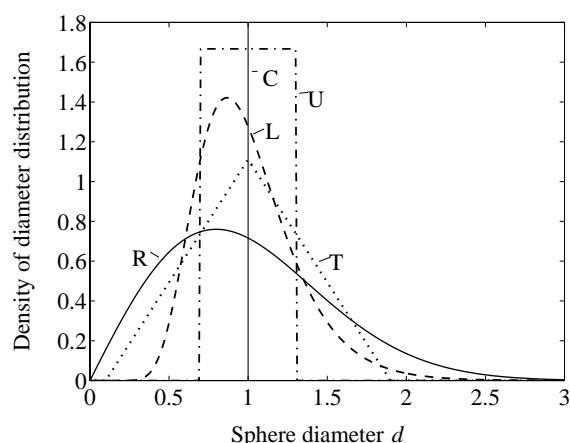


Fig. 1. *Density functions of the diameter distributions. U – uniform, T – triangular, R – Rayleigh, L – lognormal. For the parameters see text. (C – line representing constant diameters $\equiv 1$.)*

Since the relative bias is dimensionless, the results also hold for diameter distributions with the same shape but with mean \bar{d}_V different from one. To this end, B' is regarded as a function of the ratio $\varepsilon : \bar{d}_V$ or $t : \bar{d}_V$, resp.

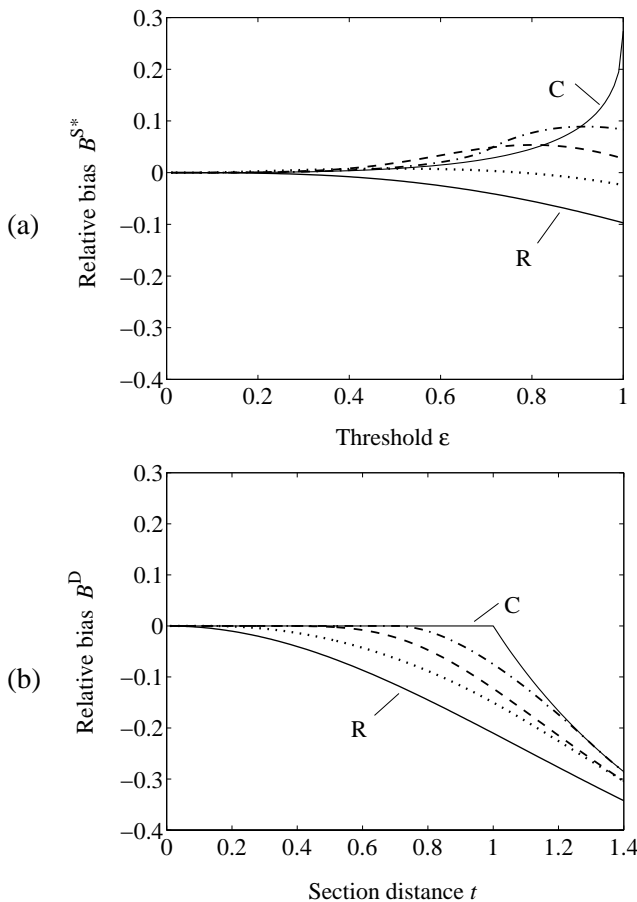


Fig. 2. Relative bias (a) of the improved Saltykov estimator and (b) of the disector as a function of the ratios $\epsilon : \bar{d}_V$ or $t : \bar{d}_V$, resp. for the diameter distributions in Fig. 1.

Fig. 2 shows (a) B^{S^*} for $\epsilon : \bar{d}_V \in [0, 1]$ and (b) B^D for $t : \bar{d}_V \in [0, 1.4]$.

The absolute value of the relative bias B^{S^*} of the improved Saltykov estimator is very small (between 0.0004 and 0.0026) for all considered distributions if $\epsilon \leq 0.2$. It still does not exceed 0.0082 for $\epsilon \leq 0.4$. Therefore a ratio “ ϵ : mean diameter” of 0.2 or even 0.4 can be recommended for practical use.

The disector is unbiased if the distance t is smaller than the smallest sphere diameter, e.g. for $t < 1$ in case of constant diameters. In general the bias of the disector depends very much on the form of the diameter distribution.

In contrast to mean or bias, the variance of the estimators also depends on the variability in spatial arrangement of the particles. In general the relations are very complex. However it is possible to give formulae for the case of independently random positioned sphere centres and independent diameters, i.e. for the Poisson case.

The variance of $\hat{N}_V^{S^*}$ is given by

$$\text{Var} \hat{N}_V^{S^*} = \frac{N_V}{A(W)} \left[\frac{4}{\pi^2} \int_{\epsilon}^{\infty} \frac{1}{d} \ln \left(\frac{d}{\epsilon} + \sqrt{\left(\frac{d}{\epsilon} \right)^2 - 1} \right) D_V(dd) + \frac{16}{\pi^2 \epsilon^2} \int_0^{\infty} (D_V(\sqrt{u^2 + \epsilon^2}) - D_V(u)) du \right], \quad (6)$$

and the variance of \hat{N}_V^D is

$$\text{Var} \hat{N}_V^D = \frac{N_V}{A(W)} \cdot \frac{1}{t} \left(1 - \frac{1}{t} \int_0^t (t-d) D_V(dd) \right), \quad (7)$$

see the Appendix.

In parallel with relative bias, relative variance $V' = \text{Var} \hat{N}_V' \cdot (\bar{d}_V A(W)) / N_V$ was calculated for both estimators as a dimensionless quantity. Fig. 3(a) shows V^{S^*} for $\epsilon \in [0, 1]$. The variance does not very much depend on the distribution of the diameters. As to be expected, it decreases with increasing ϵ . For $\epsilon = 0.2$, V^{S^*} takes values between 1.75 (constant diameters) and 2.18 (Rayleigh distribution), and for $\epsilon = 0.4$, $V^{S^*} \in [1.48, 1.71]$.

The relative variance V^D of the disector method is shown in Fig. 3(b). Obviously it depends even less on the type of diameter distribution than the variance of $\hat{N}_V^{S^*}$ does. It is quite large ($V^D > 3.25$) for $t < 0.3$, but it decreases with increasing t because then the volume of the three-dimensional counting box increases which is used for the disector estimator, i.e. the sample size increases.

PRECISION OF THE N_V ESTIMATORS

Performance of an estimator is usually expressed by the mean squared error (MSE), that is the average squared distance of the estimate to the true estimated parameter. MSE is a combination of variance and bias: $\text{MSE}' = E (\hat{N}_V' - N_V)^2 = \text{Var} \hat{N}_V' + (\text{bias} \hat{N}_V')^2$.

For any number density N_V and any window size $A(W)$, the MSE of the N_V estimators can easily be calculated from relative variance V' and relative bias B' :

$$\text{MSE}' = \frac{N_V}{\bar{d}_V A(W)} V' + N_V^2 (B')^2 = N_V^2 \left(\frac{V'}{En} + (B')^2 \right), \quad (8)$$

where $En = \bar{d}_V N_V A(W)$ is the mean number of section circles counted in the observation window or counting frame, respectively.

This enables the comparison of the precision of the estimators $\hat{N}_V^{S^*}$ and \hat{N}_V^D using the formulae and diagrams of the previous section. Clearly, MSE depends on the diameter distribution and on N_V

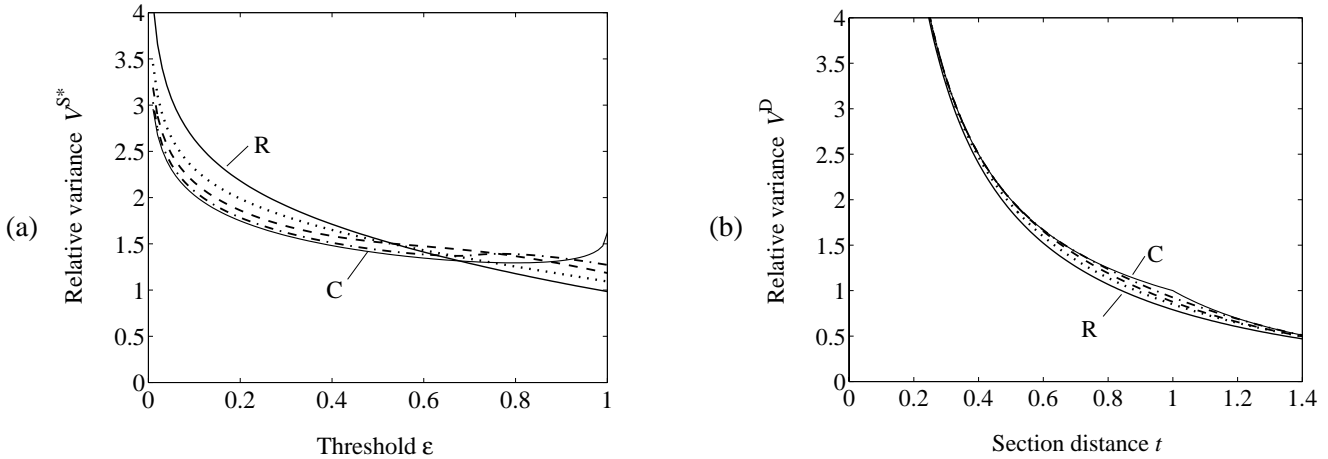


Fig. 3. Relative variance (a) of the improved Saltykov estimator and (b) of the disector as a function of the ratios $\varepsilon : \bar{d}_V$ or $t : \bar{d}_V$, resp. for the diameter distributions in Fig. 1.

itself, two factors that are intrinsic to the investigated structure and cannot be altered. Yet the experimenter can influence MSE by the choice of observation window area $A(W)$ and design parameters ε or t , resp.

A popular rule of thumb advises to chose the sample size — here represented by the window area — such that the variance part of MSE is not smaller than the bias part. This means, in terms of (8), that V^*/En should be greater or equal to $(B^*)^2$. The two parts are of equal size if the observation window or counting frame size is chosen to be $A(W)^{\text{opt}} = V^*/(\bar{d}_V N_V (B^*)^2)$.

As an example, two cases are studied of a structure with independent randomly positioned spheres of constant and independent Rayleigh distributed diameters, resp. The number density is $N_V = 1000$ [mm^{-3}], and the mean diameter is $\bar{d}_V = 0.02$ [mm]. For the disector, the section distances $t = 0.02$ and $t = 0.01$ [mm], and for the improved Saltykov estimator, $\varepsilon = 0.008$ and $\varepsilon = 0.004$ [mm] are considered, corresponding to ratios $t : \bar{d}_V$ of 1 and 0.5 resp., and $\varepsilon : \bar{d}_V$ of 0.4 and 0.2, resp.

Table 1 shows relative bias and variances, as well as the MSE for a window of area 10 (“MSE₁₀”), and the area A_{5000} of a window such that $\text{MSE} = 5000$. The absolute bias of the improved Saltykov estimator is very small for any of the considered distributions. Its contribution to MSE is practically negligible. For the physical disector, the bias strongly depends on the diameter distribution. If the structure has a relatively high proportion of small particles, a small bias can only be obtained with a small section distance t . But then the variance is very large, again leading to a large MSE. E.g. in the case of the Rayleigh distribution, \hat{N}_V^D with $t = 0.5\bar{d}_V$ has both a larger relative bias and a larger relative variance than \hat{N}_V^{S*} with $\varepsilon = 0.4\bar{d}_V$, see Table 1. As the bias of $\hat{N}_V^{S*}(0.4)$ is negligible compared to the variance and as variance of \hat{N}_V^D increases with decreasing section distance, $\hat{N}_V^{S*}(0.4)$ is better than any disector with distance $t < 0.5\bar{d}_V$. Generally spoken, the improved Saltykov estimator is particularly useful if the diameter distribution has a large range.

Taking into account that the disector uses two section planes, the total window area which is required

Table 1. $\text{MSE}_{10} = \text{MSE} [\text{mm}^{-6}]$ for a window of area 10, and $A_{5000} = \text{area} [\text{mm}^2]$ of a window such that $\text{MSE} = 5000$, of disector and improved Saltykov estimator for different parameters t and ε , resp. $N_V = 1000$ [mm^{-3}], $\bar{d}_V = 0.02$ [mm]. For further details see text above.

	constant diameters				Rayleigh distribution			
	B^*	V^*	MSE_{10}	A_{5000}	B^*	V^*	MSE_{10}	A_{5000}
$\hat{N}_V^D : t : \bar{d}_V = 1$	0	1.000	5000.0	10.0	-0.2101	0.7899	48087.9	—*
$t : \bar{d}_V = 0.5$	0	2.000	10000.0	20.0	-0.0618	1.8765	13197.6	79.2
$\hat{N}_V^{S*} : \varepsilon : \bar{d}_V = 0.4$	0.0038	1.4809	7418.5	14.9	-0.0080	1.7118	8622.3	17.3
$\varepsilon : \bar{d}_V = 0.2$	0.0004	1.7479	8739.8	17.5	-0.0010	2.1811	10906.7	21.8

—* : A_{5000} does not exist, $\text{MSE} > 5000$ for any window area.

to achieve the same precision is smaller for $\hat{N}_V^{S^*}$ even in the case of constant diameters. However measuring diameters in the case of $\hat{N}_V^{S^*}$ means acquiring more information of the window than bare counting as in the case of \hat{N}_V^D , and therefore perhaps a little higher effort to evaluate the sections.

A preliminary report was presented at the Xth International Congress for Stereology, Melbourne, Australia, 1-4 November 1999.

REFERENCES

Fullman RL (1953). Measurement of particle size in opaque bodies. *J Metals* 5:447-52.
 Howard CV, Reid S, Baddeley A, Boyde A (1985). Unbiased estimation of particle density in the tandem scanning reflected light microscope. *J Microsc* 138:203-12.
 Saltykov SA (1950). Introduction to Stereometric Metallography (Russ.) Erevan: Publ Acad Sci Armen SSR.
 Sterio DC (1984). The unbiased estimation of numbers and sizes of particles using the disector. *J Microsc* (1984) 134:127-36.
 Stoyan D, Kendall WS, Mecke J (1995). Stochastic geometry and its applications. Chichester: John Wiley & Sons.

APPENDIX: DERIVATION OF ESTIMATOR MEANS AND VARIANCES

Following the notation in Stoyan *et al.* (1995, p. 355), the particle system is represented by a spatial germ grain model (or marked point process) $\Psi_V = \{[x_n, y_n, z_n; d_n]\}$ of spherical particles, where (x_n, y_n, z_n) is the centre of the n th sphere and d_n is its diameter. The corresponding unmarked point process of particle centres has density N_V .

In order to calculate the mean of \hat{N}_V^D , we translate the number Q^- into that notation. A sphere of diameter d with centre in (x, y, z) intersects both reference and look up plane if $z - \frac{d}{2} < 0 < z + \frac{d}{2} < t$. Then it produces a section circle of radius $r = \sqrt{(d/2)^2 - z^2}$. The circle is counted in the counting frame if its centre falls inside the frame shifted by the vector (r, r) . Hence

$$\hat{N}_V^D = \frac{Q^-}{tA(W)} = \frac{1}{tA(W)} \sum_{[x,y,z;d] \in \Psi_V} \mathbf{1}_{W_{(r,r)}}(x,y) \cdot \mathbf{1}\left(z - \frac{d}{2} < 0 < z + \frac{d}{2} < t\right),$$

where $W_{(r,r)}$ represents the counting window in the reference plane shifted by the vector (r, r) and the symbol $\mathbf{1}(\cdot)$ denotes a version of the indicator function: $\mathbf{1}(A) = 1$ if the expression A is true and $\mathbf{1}(A) = 0$ if A is false.

In the following a number of expectations of sums similar to the sum above have to be calculated. Application of Campbells' theorem for marked point processes (e.g. Stoyan *et al.*, p. 108) yields for any non-negative measurable function f

$$\begin{aligned} & E\left(\sum_{[x,y,z;d] \in \Psi_V} \mathbf{1}_{W_{(r,r)}}(x,y) f(z,d)\right) \\ &= \int_0^\infty \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{W_{(r,r)}}(x,y) f(z,d) N_V dx dy dz D_V(dd) \\ &= N_V A(W) \int_0^\infty \int_{\mathbb{R}} f(z,d) dz D_V(dd) \\ &= E\left(\sum_{[x,y,z;d] \in \Psi_V} \mathbf{1}_W(x,y) f(z,d)\right). \end{aligned} \tag{9}$$

Consequently, the mean of the disector is

$$\begin{aligned} E\hat{N}_V^D &= N_V \frac{1}{t} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{W_{(r,r)}}(x,y) \cdot \\ &\quad \cdot \mathbf{1}\left(z - \frac{d}{2} < 0 < z + \frac{d}{2} < t\right) dz D_V(dd) \\ &= N_V \frac{1}{t} \left[\int_t^\infty t D_V(dd) + \int_0^t d D_V(dd) \right] \\ &= N_V \left[1 - \frac{1}{t} \int_0^t (t-d) D_V(dd) \right]. \end{aligned}$$

Under the assumption that the particles are independent randomly positioned and that their diameters are independent identically distributed, i.e. that Ψ_V is a Poisson process, the number Q^- of particles counted in the disector box is Poisson distributed. Therefore $\text{Var } Q^- = E Q^-$ and thus (7) follows from

$$\text{Var } \hat{N}_V^D = \frac{E Q^-}{t^2 A(W)^2} = \frac{E \hat{N}_V^D}{t A(W)}$$

and (5).

The improved Saltykov estimator can be written as

$$\hat{N}_V^{S^*} = \frac{2}{\pi A(W)} \sum_{[x,y,z;d] \in \Psi_V} \mathbf{1}_W(x,y) \cdot \mathbf{1}\left(-d/2 < z < d/2\right) \frac{1}{s^*(\epsilon, d, z)},$$

where

$$\begin{aligned} s^*(\epsilon, d, z) &= \mathbf{1}\left(\sqrt{d^2 - 4z^2} \geq \epsilon\right) \sqrt{d^2 - 4z^2} + \\ &\quad + \mathbf{1}\left(\sqrt{d^2 - 4z^2} < \epsilon\right) \frac{\epsilon}{2} \end{aligned}$$

is the modified section circle diameter. Following (9), the mean of $\hat{N}_V^{S^*}$ is

$$\begin{aligned} E\hat{N}_V^{S^*} &= N_V \frac{2}{\pi} \left[\int_0^\infty \int_{-\sqrt{d^2-4z^2}}^{\sqrt{d^2-4z^2}} \frac{dz}{\sqrt{d^2-4z^2}} D_V(dd) + \right. \\ &\quad \left. + \frac{2}{\varepsilon} \int_{\mathbb{R}} \int_0^\infty \mathbf{1}(4z^2 < d^2 < 4z^2 + \varepsilon^2) D_V(dd) dz \right] \\ &= N_V \frac{2}{\pi} \left[\int_\varepsilon^\infty \arccos \frac{\varepsilon}{d} D_V(dd) + \right. \\ &\quad \left. + \frac{2}{\varepsilon} \int_0^\infty \left(D_V(\sqrt{u^2 + \varepsilon^2}) - D_V(u) \right) du \right]. \end{aligned}$$

The second moment of $\hat{N}_V^{S^*}$ is

$$\begin{aligned} E \left(\frac{2}{\pi A(W)} \sum \frac{1}{s_i(\varepsilon)} \right)^2 &= \frac{4}{\pi^2 A(W)^2} E \sum \frac{1}{s_i^*(\varepsilon)^2} + \\ &\quad + \frac{4}{\pi^2 A(W)^2} E \sum_{i \neq j} \frac{1}{s_i^*(\varepsilon) s_j^*(\varepsilon)} \\ &=: V_1 + V_2. \end{aligned}$$

With the Campbell theorem (9),

$$\begin{aligned} V_1 &= \frac{4}{\pi^2 A(W)^2} E \sum_{[x,y,z;d] \in \Psi_V} \mathbf{1}_W(x,y) \mathbf{1}(|z| < \frac{d}{2}) \cdot \\ &\quad \cdot \left(\frac{\mathbf{1}(\sqrt{d^2-4z^2} \geq \varepsilon)}{d^2-4z^2} + \frac{\mathbf{1}(\sqrt{d^2-4z^2} < \varepsilon)}{\varepsilon/2} \right) \\ &= \frac{4}{\pi^2 A(W)} N_V \left[\int_{\mathbb{R}} \int_{\sqrt{4z^2+\varepsilon^2}}^\infty \frac{D_V(dd)}{d^2-4z^2} dz + \right. \\ &\quad \left. + \frac{4}{\varepsilon^2} \int_0^\infty \left(D_V(\sqrt{u^2 + \varepsilon^2}) - D_V(u) \right) du \right] \\ &= \frac{4N_V}{\pi^2 A(W)} \left[\int_\varepsilon^\infty \frac{1}{2d} \ln \frac{d + \sqrt{d^2 - \varepsilon^2}}{d - \sqrt{d^2 - \varepsilon^2}} D_V(dd) + \right. \\ &\quad \left. + \frac{4}{\varepsilon^2} \int_0^\infty \left(D_V(\sqrt{u^2 + \varepsilon^2}) - D_V(u) \right) du \right]. \end{aligned}$$

The value of V_2 depends on the distribution of Ψ_V . In the particular case of centres forming a stationary Poisson process and of independent identically distributed diameters, it is $V_2 = (E\hat{N}_V^{S^*})^2$. Then $\text{Var} \hat{N}_V^{S^*} = V_1$, which leads to (6).