S.V. Ershov, M.S. Kopylov, A.G. Voloboy

Optimal weights for bidirectional ray tracing with photon maps while mixing 3 strategies

S.V. Ershov, M.S. Kopylov, A.G. Voloboy

Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies

С.В. Ершов, М.С. Копылов, А.Г. Волобой

Оптимальные веса для случая трех стратегий в двунаправленной трассировке лучей с фотонными картами

Двунаправленная стохастическая трассировка лучей с фотонными картами — гибкий и мощный метод, однако страдающий от шума в итоговом изображении. Для его снижения используется выборка с множественной значимостью, которая объединяет результаты различных "стратегий" с весами. Оптимальные веса минимизируют шум и, таким образом, обеспечивают наилучшее качество результатов. В статье мы выводим и решаем систему интегральных уравнений, определяющих оптимальные веса; она качественно отличается от ранее исследованного случая двух стратегий. Она допускает решение в замкнутой форме как алгебраическую формулу, включающую несколько интегралов известных функций, которые могут быть вычислены при трассировке лучей.

Ключевые слова: стохастическая трассировка лучей, фотонные карты, множественная выборка по значимости, оптимальные веса.

S.V. Ershov, M.S. Kopylov, A.G. Voloboy

Optimal Weights for Bidirectional Ray Tracing with Photon Maps while Mixing 3 Strategies

Bidirectional stochastic ray tracing with photon maps is a powerful method but suffers from noise. It can be reduced by the Multiple Importance Sampling which combines results of different "strategies". The "optimal weights" minimize the noise functional thus providing the best quality of the results. In the paper we derive and solve the system of integral equations that determine the optimal weights. It has several qualitative differences from the previously investigated case of mixing two strategies, but further increase of their number beyond 3 retains the qualitative features of the system. It can be solved in a closed form i.e. as an algebraic formula that include several integrals of the known functions that can be calculated in ray tracing.

Key words: stochastic ray tracing, photon maps, multiple importance sampling, optimal weights.

Abbreviations and notations

MCRT = Monte-Carlo ray tracing

FMCRT = forward Monte-Carlo ray tracing. It is tracing of rays from light sources toward scene objects accumulating of illumination on scene objects. Usually it is used to calculate the secondary (indirect) illumination.

BMCRT = backward Monte-Carlo ray tracing. It is tracing of rays from camera through virtual screen toward scene objects. Usually it is used to get the virtual scene image.

BDF = bi-directional scattering function. It describes surface luminance as a function of the illumination and observation directions. In this paper $f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{x})$ is BDF (in luminance units) of the surface point $\boldsymbol{x}$ for illumination direction $\boldsymbol{v}$ and viewing direction $\boldsymbol{u}$.

BDD = backward diffuse depth. It is a specific parameter of a hybrid ray tracing, when FMCRT calculates illumination and BMCRT is used to convert it to the observed luminance. In this method the backward ray usually has a limited "length" and terminates after BDD diffuse events.

$(\boldsymbol{x}_0, \boldsymbol{x}_1, ...)$ light (or forward) path. It starts from the light source position $\boldsymbol{x}_0$. In each vertex it can be scattered or absorbed, but keeps unit "energy".

$(\boldsymbol{y}_{-1}, \boldsymbol{y}_0, \boldsymbol{y}_1, ...)$ camera (or backward) path. It starts from the camera origin $\boldsymbol{y}_{-1}$ which is the same for all pixels and all rays through this pixel. $\boldsymbol{y}_0$ is the first hit which depends on pixel but is the same for all rays for this pixel. In each vertex the ray is only scattered (never absorbed!) and its "energy" decreases

$(\boldsymbol{z}_0, \boldsymbol{z}_1, ...)$ joint path in which $\boldsymbol{z}_i$ can be taken from either the camera or the light subpath depending on the intersection type. The camera origin and light source position are dropped as always fixed.

$E(\boldsymbol{y}_0, ..., \boldsymbol{y}_i)$ camera ray energy before hitting $\boldsymbol{y}_{i+1}$.

$\boldsymbol{n}(\boldsymbol{x})$ local normal at surface point $\boldsymbol{x}$.

$L(\boldsymbol{x}_b, \boldsymbol{x}_a)$ is the full luminance at the point $\boldsymbol{x}_a$ in direction $\overrightarrow{\boldsymbol{x}_a \boldsymbol{x}_b}$. Similarly, $L_0(\boldsymbol{x}_b, \boldsymbol{x}_a)$ is the direct luminance and $L_d(\boldsymbol{x}_b, \boldsymbol{x}_a)$ is the diffuse luminance.

R radius of the "integration sphere" (distance threshold below which the light and camera vertices can be merged).

$S = \pi R^2$ area of the "integration sphere" (this is a 2D sphere in the tangent plane, i.e. rather a circle).

N_F the number of light rays traced during one iteration.

N_B the number of camera rays traced *through this pixel* during one iteration.

1 Introduction

Nowadays lighting simulation is widely used in realistic computer graphics and for designing new materials and optical systems. This group of methods mainly includes the simulation of light transport using the Metropolis method [1] and stochastic ray tracing [2] and their bidirectional modifications. Among them, we consider the bidirectional Monte-Carlo ray tracing with photon maps (BDPM) [3, 4]. The main problem of all stochastic methods is noisy result. Therefore, the noise reduction problem is always important, and it is considered in many works, e.g., [5, 6, 7]. The level of noise in BDPM mainly depends on the random scattering of the forward and backward rays, on the choice of the vertex for their merging and, finally, on the number of forward and backward rays traced in one iteration step. The majority of studies proposed to decrease the noise by application of the multiple importance sampling (MIS) technique [5, 6, 8].

An important foundation here is the Veach's work [9]. Veach results are based on the theorem which assumes independent samples. But direct application of them to BDPM would be incorrect. The samples (i.e. full paths connecting the source and the camera) happen to be not independent because in BDPM the same light path is merged with many camera paths and vice versa. So the resulting joint full paths have common part and thus they are not independent. In [10] the problem of the optimal weights in a limited MIS for mixing contributions from *two* first camera vertices was considered. This limited case already explains the main idea of calculation of the optimal weight for the strategies in BDPM, i.e. that unlike the famous Veach heuristic for the optimal weight in MCRT [9], [4, 11] where they obey a system of algebraic equations, now they obey the system of integral equations.

However, mixing just two strategies has several deficiencies. For example, let us suppose that the BDF at the first two camera vertices is sharp, while the 3rd vertex has a smooth BDF. Mixing of just the two strategies (gathering illumination at the 1st and st the 2nd vertices, respectively) may choose only among them. But gathering illumination at a point where BDF is sharp is *bad* (has high noise). This can be seen from the formulae for the optimal weights from [10]: the sharper the BDF the lower the weight for the corresponding vertex. Therefore, mixing two strategies allows only to choose less evil among

the two. Meanwhile the gathering illumination at the 3rd vertex with smooth BDF will be good. Or maybe at some another (later) vertex. Therefore it is of interest to generalize the problem to mixing more than 2 strategies.

The first step, and it is principal, is mixing of three strategies. This case in investigated in the current work, and we shall see that it is qualitatively different from mixing two strategies. The main reason is that while for mixing two strategies we had only one variable weight, now we have two *different families* of weights, each with its own normalization conditions. The first consists of 2 weights and one of them is dependent. The second family consists of three weights, any two of them can be considered as independent and the remaining one is dependent on them. In case of mixing four and more strategies the situation already remains qualitatively the same, just there will be three and more families of weights. All independent weights from all families are coupled in the common system of integral equations and this holds if mixing more than 3 strategies. Meanwhile when mixing 2 strategies we had just one weight and, correspondingly, one integral equation.

Later we shall see that the resulting system of integral equations admits a "local" approximation when it reduces to the system of algebraic equations because the integral terms can be neglected. For two strategies the formula relates some values at the two vertices of which both are the "end points". Therefore one can not deduce what it will turn into if adding another vertex or more of them. Meanwhile this formula for 3 vertices already admits an "intuitive induction" and one can imagine what it must turn into for mixing 4, 5 and more strategies.

Therefore, investigation of mixing of three strategies (which correspond to gathering illumination at the 1st, 2nd and the 3rd vertices of the camera path) is a necessary next step. It is performed in the current paper where we shall derive the system of equations the optimal weights obey.

2   Related works

A good overview of the studies performed by the Krivanek's group [12] had been recently published. It explains a wide variety of different methods, including bi-directional path tracing, light transport etc. providing a concise explanation of the underlying ideas followed by a list of detailed publications. We dare recommend this course to all engaged in related areas.

There is a part just related to MIS in bi-directional path tracing (pp. 214-246 of [12]) and a detailed paper[13]. They improve the classic Veach results [9]. There are two different formulae to calculate the optimal weights by Veach, named the "balance heuristic" and the "power heuristic", and this is because Veach minimized not the variance itself but rather its bounds. Veach theorem

states that the true minimum noise will not deviate from this estimate much giving the estimation of that difference. Nevertheless, the true optimization is undoubtedly better. The resulting formulae (13) for the optimal weights in [13] consist of the parts present in the "balance heuristic", but with scale coefficients α are calculated from the linear system, see eqs. (10)–(12) in [13].

The general scheme of calculation of the optimal weights in [13] and in our papers [10, 14] is rather similar. In both approaches the noise (variance) is a quadratic integral functional on the weights. Its minimum is then achieved for the weights which obey a system of linear integral equations (eq. (28) in [13] vs our (29), (30)). Their α_i is somewhat similar to our $G_i(\boldsymbol{x}_1)$ though the latter depends on the space point (i.e. the vertex of the joint path). Then the weights themselves are written as the linear combination of the integral terms and then integrated which gives a system in that integral terms (eq. (32) from [13] vs our eq. (38)). Our procedure though does not end here because the integrals $G_m(\boldsymbol{x}_1)$ are space functions unlike their α_i and so we need yet another iteration of a similar procedure. Eventually the optimal weights in both their and our investigations are expressed as algebraic functions with coefficients being integrals (averages!) of known scene functions. They can be calculated by Monte-Carlo averaging.

So there solution is also closer to ours than the algebraic "local" formulae of the balance or power heuristics. In spite of the "overall similarity", our formulae are different. For example, our weights depend on scene luminance at the joint path vertex, while in [13] they do not. Also, our weights depend on the number of camera rays per pixel N_B and the number of light rays N_F, common to all pixels.

The difference is because these weights are for two *different* variants of bi-directional MCRT. In [13] it is the *bi-directional path tracing* (BDPT) while we use the bi-directional ray tracing *with photon maps* (BDPM). Roughly, they trace *one* light path of the desired "depth" (the number of segment) and *one* camera path with the desired depth. These depths determine the "strategy" (so it is decided *a priori*, as is in the classic MIS). After that the ends of these sub-paths are *connected* with a straight segment, and if it is not occluded, then we get the joint path and increment the accumulated pixel luminance by its "importance function". If the segment is occluded, the increment is 0. After that the pair of light and camera paths is discarded and a new one is generated.

In BDPM, the process is basically different. We trace N_B camera rays and N_F light rays, then *each* camera path is checked against *each* light path, and if they have close vertices, these are *identified* and we get the joint path and increment the accumulated pixel luminance by its "importance function". If their is no close vertices, the increment is 0.

As a result, the BDPT operates *independent* "samples" (joint paths), while

BDPM operates dependent *"samples"*. Indeed, we merge the same camera sub-path with different light paths, and thus can obtain several joint paths which have the same "camera" tail and thus are not independent. Also, in BDPT which follows the classic MIS approach, the "strategy" is chosen before generating the random sample, and in principle one can decide freely and arbitrary how many samples to generate with which strategy. In BDPM the situation is different. We do not know in advance which vertices of the light and camera path will be close (or none), so the strategy is determined *a posteriori* and depends on the sample (path). The number of samples for each strategy is thus not controllable directly. It is even unknown in advance.

3 Bi-directional Ray Tracing with Photon Maps

*Here and below all calculations are for **one** pixel.*

For the sake of simplicity, the total flux of all scene lights is assumed 1 to not bother about scaling between the density of photons and irradiance. We also assume that the light source is point one and that it is unique in the scene.

It is well known that the luminance of a pixel can be written as an average over the paths connecting light source and camera which average can be naturally calculated with a sort of Monte-Carlo integration [9]. The problem is efficiency. The probability that a forward or *light* ray hits the (usually tiny) camera pupil is very low. Similarly, the probability that a backward or *camera* ray hits a point or line light source is exactly 0.

A natural remedy is that the "light source end" is generated with the forward MCRT (*from* light source) while the "camera end" is generated by the backward MCRT (*from* camera) and then these two "half-paths" are *somehow* joined. This "somehow" is the crucial point and there is a real lot of different methods, mainly vertex *connection*, vertex *merging* and *photon maps*. All of them have their strong and weak points, none being a magic wand. We shall investigate the photon maps or reverse photon maps [15, 16] which is the same from noise estimation point of view.

3.1 Base idea of BDPM

The idea of joining sub-paths in the reverse photon map method is very simple and natural. Roughly, first we do BMCRT and its hit points are stored in the map. Then the FMCRT starts. Whenever its ray hits a surface, we ask from the reverse photon map all camera hits which are close to this point (i.e. the forward ray hit is within the integration sphere about the camera hit). If it is
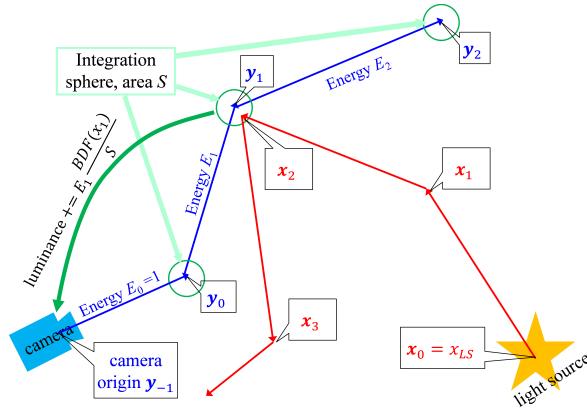
Figure 1: The main idea of BDPM: we trace camera path (blue) and light path (red) and when a vertex of the latter is close to a camera vertex, we can joint the paths neglecting this small distance. After that the accumulated luminance of the pixel related to the camera sub-path is incremented.

so, we unify these close vertices and thus obtain the joint path connecting the camera and the light source (Fig. 1).

For each this joint path, we increment of luminance of pixel corresponding to camera ray. This increment is the *contribution* of the joint path. For the correct estimation of luminance this contribution must equal to the ratio of the *target* probability density for this joint path (or its "importance", [9]) to the probability density of its generation in the above BDPM procedure. So the contribution $\mathcal{C}$ equals

$$\mathcal{C} = \frac{E_i \mathrm{BDF}(\boldsymbol{y}_i)}{S}$$

where $\boldsymbol{y}_i$ is the camera hit point where the sub-paths joint, E_i is the energy of the camera ray *before* this hit (i.e. *after* the previous hit), $\mathrm{BDF}(\boldsymbol{x})$ is the surface BDF at point $\boldsymbol{x}$ and S is the area of the integration sphere (so that S^{-1} gives the estimation of the irradiance from *single* incident ray with unit energy). The whole procedure is depicted in Fig. 1.

A bit more formally, if the n-th a vertex $\boldsymbol{x}_n$ of the light path $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \cdots \to \boldsymbol{x}_n \to \cdots$ (here $\boldsymbol{x}_0$ is the light source origin position) is near the i-th vertex $\boldsymbol{y}_i$ of the camera path $\boldsymbol{y}_{-1} \to \boldsymbol{y}_0 \to \cdots \to \boldsymbol{y}_i \to \cdots$ (here $\boldsymbol{y}_{-1}$ is the camera origin where the rays for all pixels starts from), then the full path is

$$\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \cdots \to \boldsymbol{x}_{n-1} \to \boldsymbol{y}_i \to \cdots \to \boldsymbol{y}_1 \to \boldsymbol{y}_0 \to \boldsymbol{y}_{-1},$$

its *contribution* (to the pixel luminance) being

$$\mathcal{C} = S^{-1} E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1}) f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_i\boldsymbol{y}_{i-1}}, \boldsymbol{y}_i)$$
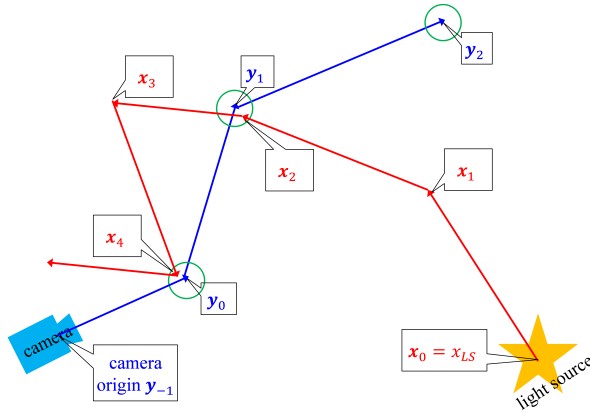
Figure 2: Camera and light subpaths can "intersect" at different vertices. In this example there are two close pairs of vertices: $\boldsymbol{x}_4 \approx \boldsymbol{y}_0$ and $\boldsymbol{x}_2 \approx \boldsymbol{y}_1$ which produce two different joint paths $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \boldsymbol{x}_2 \to \boldsymbol{x}_3 \to \boldsymbol{y}_0 \to \boldsymbol{y}_{-1}$ and $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \boldsymbol{y}_1 \to \boldsymbol{y}_0 \to \boldsymbol{y}_{-1}$ of different length.

where $\overrightarrow{\boldsymbol{ab}}$ is the *unit* vector from point $\boldsymbol{a}$ to point $\boldsymbol{b}$, $f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{x})$ is BDF (in luminance units) of the surface point $\boldsymbol{x}$ for illumination direction $\boldsymbol{v}$ and viewing direction $\boldsymbol{u}$, and $E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1})$ is the energy of the camera ray *before hitting* $\boldsymbol{y}_i$. This energy (or transmission factor in [4] terms) is defined as usual: it is 1 just after leaving the camera, i.e. $E(\boldsymbol{y}_{-1}) = 1$ and then

$$
\begin{aligned}
E(\boldsymbol{x}_0) &= \mu(\overrightarrow{\boldsymbol{x}_{-1}\boldsymbol{x}_0}, \boldsymbol{x}_0), \\
E(\boldsymbol{x}_0, ..., \boldsymbol{x}_m) &= \mu(\overrightarrow{\boldsymbol{x}_{m-1}\boldsymbol{x}_m}, \boldsymbol{x}_m) E(\boldsymbol{x}_0, ..., \boldsymbol{x}_{m-1}), \qquad m = 1, 2, ... \quad (1)
\end{aligned}
$$

where

$$
\mu(\boldsymbol{u}, \boldsymbol{x}) \equiv \int f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{x}) \left| (\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{x})) \right| d^2 \boldsymbol{v}, \qquad (2)
$$

$\boldsymbol{n}(\boldsymbol{x})$ being the local normal at the point $\boldsymbol{x}$.

3.2   Final form of the path contribution and the gathering equation

In principle, the camera and light sub-paths can have several "intersections" (although this is rare in case of small integration area) as shown in Fig. 2

The increment of the accumulated pixel luminance in the above example is the sum of contributions of these two joint paths. Naturally this full increment is just the *contribution of all* joint *paths obtained from all intersections* of the given camera and light sub-paths.

The general formula that allows an arbitrary number of "intersections" (close pairs of vertices) calculates the contribution from the camera path

$(\boldsymbol{y}_{-1}, \boldsymbol{y}_0, \boldsymbol{y}_1, ...)$ and light path $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, ...)$ as the weighted sum

$$
\begin{aligned}
C \;=\; & \sum_{i \geq 0} \sum_{n \geq 1} K(\boldsymbol{y}_i - \boldsymbol{x}_n) w_{i+n,i}(\boldsymbol{y}_0, \boldsymbol{y}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n-1}, ..., \boldsymbol{x}_1) \\
& \times E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1}) f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_i\boldsymbol{y}_{i-1}}, \boldsymbol{y}_i)
\end{aligned}
$$

over *all pairs* from which the merging kernel $K(r)$ (which vanishes outside the integration sphere, i.e. for $r > R$) selects the close ones. The weight $w_{l,i} \leq 1$ where l is the full "length" (the number of vertices) of the joint path and $0 \leq i \leq l - 1$ is the index of the camera vertex where the subpaths merge can depend on the joint path. We dropped the *fixed* points $\boldsymbol{x}_0$ and $\boldsymbol{y}_{-1}$ from the arguments of the weight. These weights are needed to avoid multiple counting because the same joint path can be obtained by different "strategies" i. For this purpose

$$
\sum_{i=0}^{l-1} w_{l,i} = 1 \tag{3}
$$

for each l.

REMARK. The averaging over the full paths is a sort of MCRT, though a bit special since now different paths may be not statistically independent (which is natural since two different full paths may have the same camera sub-path and thus be related through it). This dependence of *samples* affects the variance but not the expectation. Therefore the theorems from [9] still hold which state that (3) is sufficient for the mathematical expectation of the calculated pixel luminance be correct. For our particular situation this is derived in Appendix B.

The term with $n = 1$ is for the *direct illumination* because the light ray had not yet scattered before hitting the integration sphere. Meanwhile, usually in BDPM the direct illumination is taken not from the photon maps but "deterministically". In this case the term with $n = 1$ must be replaced with

$$
w_{i+n,i}(\boldsymbol{y}_0, \boldsymbol{y}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_0) E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1}) L_0(\boldsymbol{y}_{i-1}, \boldsymbol{y}_i)
$$

where $L_0(\boldsymbol{y}_{i-1}, \boldsymbol{y}_i)$ is the direct luminance at the point $\boldsymbol{y}_i$ in direction $\overrightarrow{\boldsymbol{y}_i\boldsymbol{y}_{i-1}}$. The contribution from the camera path $(\boldsymbol{y}_{-1}, \boldsymbol{y}_0, \boldsymbol{y}_1, ...)$ and light path $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, ...)$ then becomes

$$
\begin{aligned}
C \;=\; & \sum_{i \geq 0} w_{i+n,i}(\boldsymbol{y}_0, \boldsymbol{y}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_0) E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1}) L_0(\boldsymbol{y}_{i-1}, \boldsymbol{y}_i) \\
& + \sum_{i \geq 0} \sum_{n \geq 2} K(\boldsymbol{y}_i - \boldsymbol{x}_n) w_{i+n,i}(\boldsymbol{y}_0, \boldsymbol{y}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n-1}, ..., \boldsymbol{x}_1) \\
& \times E(\boldsymbol{y}_0, ..., \boldsymbol{y}_{i-1}) f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_i\boldsymbol{y}_{i-1}}, \boldsymbol{y}_i)
\end{aligned}
$$

The pixel luminance is then the average of that contribution over both the camera and light rays. The contribution C is a function of the camera and light paths.

Notice that the above formula calculates the *estimator of the pixel luminance* (= the value whose average gives that luminance) as the sum over all "intersections" at different camera vertices with corresponding weights that depend on the index of that vertex. This is nothing but the Multiple Importance Sampling (MIS for short) because it sums over all possible joint paths i.e. the "samples" in the path averaging.

3.3 Partial MIS-3

The full MIS mixes intersections at *all* vertices of the given camera path and thus includes an infinite number of weights. In spite of several obvious advantages, it is expensive for the practical implementation, because it requires that we trace camera rays to *infinite depth* and then store all their numerous vertices in the (reverse) photon maps. Also it includes an infinite number of weights which, being the functions of the joint path, may depend on an *arbitrary large* number of arguments.

Therefore it makes sense to work with a limited method that operates intersections at the few first camera vertices. Also, the weights in it can depend only on the few vertices of the joint path that are closer to camera, so they are

$$w_{l,i}(\boldsymbol{z}_0, \boldsymbol{z}_1, ..., \boldsymbol{z}_i), \qquad i = 0, ..., M, \qquad l > i$$

where $\boldsymbol{z}_j$ are the vertices of the joint path (numbered from camera end!) which depending from the intersection type can be taken from either the camera or the light subpath. We drop $\boldsymbol{x}_0$ (light source position) and $\boldsymbol{y}_{-1}$ (camera origin) from this set because they are always the same, fixed points.

In this formulation the weights with large first index $l > M$ all depend on the same number of arguments. Therefore it is possible impose the limitation that they are all the same function:

$$w_{l,i}(\boldsymbol{z}_0, \boldsymbol{z}_1, ..., \boldsymbol{z}_M) = w_{M,i}(\boldsymbol{z}_0, \boldsymbol{z}_1, ..., \boldsymbol{z}_M), \qquad l > M, \qquad i = 0, ..., M$$

This means that the weights do *not* depend on the starting part of the light path neither through the weight's index nor through its arguments (the points).

Notice that these restrictions of the weights are compliant with the general requirements on the weights in MIS [9] which require the weights being deterministic functions of the joint path (but may be effectively independent of its "tail"!) and that the sum of weight in each "family" be 1.
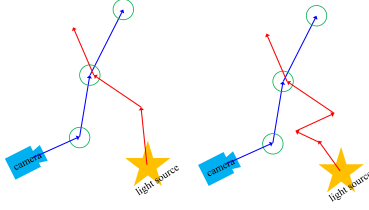
Figure 3: The weight is the same in the right panel in spite of the longer initial part of light path.

In current paper we shall consider the case when we mix contributions from *three* first camera vertices (MIS-3). Now the *gathering equation* becomes

$$
\begin{aligned}
C \;=\; & L_0(\boldsymbol{y}_{-1}, \boldsymbol{y}_0) w_{1,0}(\boldsymbol{y}_0) + E(\boldsymbol{y}_0) L_0(\boldsymbol{y}_0, \boldsymbol{y}_1) w_{2,1}(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
& + E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_1, \boldsymbol{y}_2) w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) \\
& + K(\boldsymbol{y}_0 - \boldsymbol{x}_2) w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
& + \sum_{n \geq 3} K(\boldsymbol{y}_0 - \boldsymbol{x}_n) w_{3,0}(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2}) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
& + \sum_{n \geq 2} K(\boldsymbol{y}_1 - \boldsymbol{x}_n) w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_{n-1}) E(\boldsymbol{y}_0) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_1 \boldsymbol{y}_0}, \boldsymbol{y}_1) \\
& + \sum_{n \geq 2} K(\boldsymbol{y}_2 - \boldsymbol{x}_n) w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_2 \boldsymbol{y}_1}, \boldsymbol{y}_2) \quad (4)
\end{aligned}
$$

which is an obvious generalization of the case of camera path with just 2 vertices [10]. The contribution C is a function of the camera and light paths.

Which weights relate to which intersection is depicted in Fig. 4.

These weights obey the usual normalization conditions:

$$
\begin{aligned}
w_{1,0}(\boldsymbol{z}_0) &= 1 \\
w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) + w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) &= 1 \quad \forall \boldsymbol{z}_1 \\
w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &= 1 \quad \forall \boldsymbol{z}_1, \boldsymbol{z}_2
\end{aligned}
$$

which, as we prove in **Appendix** B, are sufficient to provide correct average of pixel luminance.

4 Noise

In BDPM (with or without weights) the variance of the pixel luminance calculated from N_F forward rays and N_B backward rays (started from the same pixel) obeys the general law [7]:
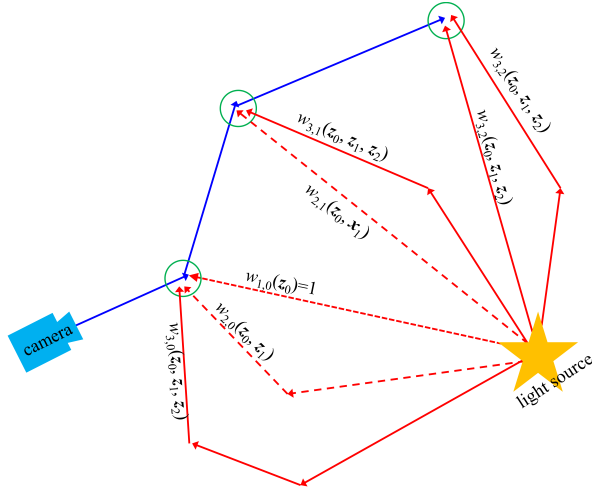
Figure 4: Weights for different intersections in BDPM. Here $\boldsymbol{z}_{-1}$, $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ etc denote the vertices of the *joint* path counted from camera (regardless of which subpath they are from: camera or light). Obviously, always $\boldsymbol{z}_{-1} = \boldsymbol{y}_{-1}$ is the camera origin and $\boldsymbol{z}_0 = \boldsymbol{y}_0$ is the first hit which is fixed for the given pixel.

$$V = \frac{1}{N_F N_B}\left(\langle\langle C^2\rangle\rangle - \langle\langle C\rangle\rangle^2\right) + \frac{1 - N_F^{-1}}{N_B}\left(\langle\langle C\rangle_F^2\rangle_B - \langle\langle C\rangle\rangle^2\right)$$
$$+ \frac{1 - N_B^{-1}}{N_F}\left(\langle\langle C\rangle_B^2\rangle_F - \langle\langle C\rangle\rangle^2\right) \tag{5}$$

Here $\langle\cdot\rangle_B$ is the averaging over the BMCRT ensemble for the fixed FMCRT ray and $\langle\cdot\rangle_F$ is the averaging over the FMCRT ensemble for the fixed camera ray. Notice the linear term $\langle\langle C\rangle\rangle$ is independent from the order of averaging so we drop subscripts here. It is also independent from weights, while $\langle\langle C^2\rangle\rangle$ and $\langle\langle C\rangle_F^2\rangle_B$ depend on them.

Averaging over the ensemble of light paths resp. camera paths is

$$\langle\cdot\rangle_F = \int (\cdot)p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n, ...)d\boldsymbol{x}_1 \cdots d\boldsymbol{x}_n \cdots \tag{6}$$

$$\langle\cdot\rangle_B = \int (\cdot)p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2)d\boldsymbol{y}_1 d\boldsymbol{y}_2 \tag{7}$$

where p_F and p_B are the probability densities of the light and camera paths (the latter for fixed given pixel). Since we assume that FMCRT uses Russian roulette to kill rays while keeps ray energy, p_F is not normalized. These densities can be calculated from the recurrence relations similar to that in [4], see **Appendix A**. The *fixed* points $\boldsymbol{x}_0$ and $\boldsymbol{y}_1$ are not included in the averaging. The point $\boldsymbol{y}_0$ depends on pixel, but for given pixel it is also fixed and thus is not also included in the averaging.

Below we shall sometimes use the spatial and sometimes angular probability densities keeping in mind the obvious relation between differentials

$$
\begin{aligned}
d^2(\overrightarrow{\boldsymbol{x}\boldsymbol{y}}) &= s(\boldsymbol{x}, \boldsymbol{y})d\boldsymbol{y} \\
s(\overrightarrow{\boldsymbol{x}\boldsymbol{y}}) &\equiv \frac{\left|(\overrightarrow{\boldsymbol{x}\boldsymbol{y}} \cdot \boldsymbol{n}(\boldsymbol{y}))\right|}{\left|\boldsymbol{x} - \boldsymbol{y}\right|^2}
\end{aligned}
\tag{8}
$$

where $\boldsymbol{n}(\boldsymbol{y})$ is the local normal at the point $\boldsymbol{y}$ and $\overrightarrow{\boldsymbol{x}\boldsymbol{y}}$ is the unit vector from $\boldsymbol{x}$ to $\boldsymbol{y}$.

Notice that when $N_F \to \infty$ while N_B is fixed the noise does not vanish. This remaining noise $N_B^{-1}\left(\langle\langle C\rangle_F^2\rangle_B - \langle\langle C\rangle\rangle^2\right)$ can be naturally termed the BMCRT noise, and $\langle\langle C\rangle_F^2\rangle_B$ named the BMCRT term. Similarly, when $N_B \to \infty$ while N_F is fixed, the remaining noise $N_F^{-1}\left(\langle\langle C\rangle_B^2\rangle_F - \langle\langle C\rangle\rangle^2\right)$ is termed the FMCRT noise and $\langle\langle C\rangle_B^2\rangle_F$ is named the FMCRT term. The last quadratic average $\langle\langle C^2\rangle\rangle$ will be named the cross term.

Now let us calculate the weight-dependent quadratic averages $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle\langle C\rangle_F^2\rangle_B$ for the case of the simplest integration kernel K which is the indicator function

$$
K(r) = \begin{cases} \frac{1}{\pi R^2}, & r \le R \\ 0, & r > R \end{cases}
$$

The area of the integration area is denoted $S = \pi R^2$ and *it is assumed small.*

4.1 Cross term

The square of C given by (4) contains

1) Products of kernel K and a smooth function. Since the kernel is weakly close to the delta-function, the average of this term is $O(1)$

2) Squares of kernel times a smooth function. Since our kernel is the indicator function and thus the average of that term is $S^{-1}O(1)$, see above

3) Product of kernels with different arguments times a continuous function. The product of kernels of different arguments can also reach $S^{-1}K$ but only in the area where the two kernels overlap. Their arguments are random and the probability of intersection is $O(S)$, so the average of the product of kernels with different arguments is $O(1)$, which is formally proved in **Appendix** C.

Therefore for calculations of averages of C^2 for small S we can retain in C^2 only the squares of kernels:

$$C^2 = \frac{1}{S}K(\boldsymbol{y}_0 - \boldsymbol{x}_2)w_{2,0}^2(\boldsymbol{y}_0, \boldsymbol{x}_1)f^2(\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0}, \boldsymbol{y}_0)$$

$$+ \frac{1}{S}\sum_{n\geq3}K(\boldsymbol{y}_0 - \boldsymbol{x}_n)w_{3,0}^2(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2})f^2(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0}, \boldsymbol{y}_0)$$

$$+ \frac{1}{S}\sum_{n\geq2}K(\boldsymbol{y}_1 - \boldsymbol{x}_n)w_{3,1}^2(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_{n-1})E^2(\boldsymbol{y}_0)f^2(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1}, \boldsymbol{y}_1)$$

$$+ \frac{1}{S}\sum_{n\geq2}K(\boldsymbol{y}_2 - \boldsymbol{x}_n)w_{3,2}^2(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2)E^2(\boldsymbol{y}_0, \boldsymbol{y}_1)f^2(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \boldsymbol{y}_2)$$

$$+ (\cdots) \tag{9}$$

while the omitted terms $(\cdots)$ are inessential since have much smaller average: $\langle\langle\cdots\rangle\rangle = O(1)$.

Now we must average this approximation to C^2 over the FMCRT ensemble, i.e. $p_F(\boldsymbol{x}_1, \boldsymbol{x}_2, ...)d\boldsymbol{x}_1 d\boldsymbol{x}_2 \cdots$. It can be done very similarly to the derivation of (47) in **Appendix** B. Indeed, the above expression for C^2 differs from the 4 last lines in (9) averaged there only in that weights, BDFs and energies are now squared. Therefore applying the same derivation we arrive at

$$\langle C^2\rangle_F \approx \frac{1}{S}\int \left(w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1)f(\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0}, \boldsymbol{y}_0)\right)^2 L_0(\boldsymbol{y}_0, \boldsymbol{x}_1)$$

$$\times \left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0})\right| s(\boldsymbol{y}_0, \boldsymbol{x}_1)d\boldsymbol{x}_1$$

$$+ \frac{1}{S}\int \left(w_{3,0}(\boldsymbol{y}_0, \boldsymbol{x}_1, \boldsymbol{x}_2)f(\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0}, \boldsymbol{y}_0)\right)^2 \left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0})\right| L(\boldsymbol{x}_1, \boldsymbol{x}_2)$$

$$\times f(\overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1}, \overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0}, \boldsymbol{x}_1)\left|(\boldsymbol{n}(\boldsymbol{x}_1)\cdot\overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1})\right| s(\boldsymbol{y}_0, \boldsymbol{x}_1)s(\boldsymbol{x}_1, \boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2$$

$$+ \frac{1}{S}E^2(\boldsymbol{y}_0)\int \left(w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_2)f(\overrightarrow{\boldsymbol{x}_2\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1}, \boldsymbol{y}_1)\right)^2 L(\boldsymbol{y}_1, \boldsymbol{y}_2)$$

$$\times \left|(\boldsymbol{n}(\boldsymbol{y}_1)\cdot\overrightarrow{\boldsymbol{x}_2\boldsymbol{y}_1})\right| s(\boldsymbol{y}_1, \boldsymbol{x}_2)d\boldsymbol{x}_2$$

$$+ \frac{1}{S}w_{3,2}^2(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2)E^2(\boldsymbol{y}_0, \boldsymbol{y}_1)$$

$$\times \int f^2(\overrightarrow{\boldsymbol{x}_3\boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \boldsymbol{y}_2)L(\boldsymbol{y}_2, \boldsymbol{x}_3)\left|(\boldsymbol{n}(\boldsymbol{y}_2)\cdot\overrightarrow{\boldsymbol{x}_3\boldsymbol{y}_2})\right| s(\boldsymbol{y}_2, \boldsymbol{x}_3)d\boldsymbol{x}_3$$

$$+ O(1)$$

The last integral for which we introduce notation

$$b(\boldsymbol{x}_1, \boldsymbol{x}_2) \equiv \int f^2(\overrightarrow{\boldsymbol{x}_3\boldsymbol{x}_2}, \overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1}, \boldsymbol{x}_2)L(\boldsymbol{x}_2, \boldsymbol{x}_3)\left|(\boldsymbol{n}(\boldsymbol{x}_2)\cdot\overrightarrow{\boldsymbol{x}_3\boldsymbol{x}_2})\right| s(\boldsymbol{x}_2, \boldsymbol{x}_3)d\boldsymbol{x}_3$$

is independent of weights and is much similar to the integral which gives diffuse luminance of $\boldsymbol{x}_2$ towards $\boldsymbol{x}_1$

$$L_d(\boldsymbol{x}_1, \boldsymbol{x}_2) = \int f(\overrightarrow{\boldsymbol{x}_3\boldsymbol{x}_2}, \overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1}, \boldsymbol{x}_2) L(\boldsymbol{x}_2, \boldsymbol{x}_3) \left| (\boldsymbol{n}(\boldsymbol{x}_2) \cdot \overrightarrow{\boldsymbol{x}_3\boldsymbol{x}_2}) \right| s(\boldsymbol{x}_2, \boldsymbol{x}_3) d\boldsymbol{x}_3$$

Now we must average this $\langle C^2 \rangle_F$ over the BMCRT ensemble i.e. over $p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) d\boldsymbol{y}_1 d\boldsymbol{y}_2$. Substituting $p_B(\boldsymbol{y}_0, \boldsymbol{y}_1)$ and $p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2)$ from (40) one finally arrives at

$$
\begin{aligned}
\langle\langle C^2 \rangle\rangle \approx{} & S^{-1} \int w_{2,0}^2(\boldsymbol{z}_0, \boldsymbol{z}_1) f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) \\
& \times \left| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}) \right| s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1 \\
& + S^{-1} \int w_{3,0}^2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& + S^{-1} \int w_{3,1}^2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E(\boldsymbol{z}_0) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& + S^{-1} \int w_{3,2}^2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{b(\boldsymbol{z}_1, \boldsymbol{z}_2)}{L(\boldsymbol{z}_1, \boldsymbol{z}_2)} \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& + O(1) \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (10)
\end{aligned}
$$

where $\boldsymbol{z}_{-1}$, $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ denote the vertices of the *joint* path counted from camera (regardless of which subpath they are from: camera or light). Obviously, always $\boldsymbol{z}_{-1} = \boldsymbol{y}_{-1}$ is the camera origin and $\boldsymbol{z}_0 = \boldsymbol{y}_0$ is the first hit which is fixed for the given pixel. The function ρ is defined as

$$
\begin{aligned}
\rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \equiv{} & L(\boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \quad (11) \\
& \times (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}) s(\boldsymbol{z}_0, \boldsymbol{z}_1) s(\boldsymbol{z}_1, \boldsymbol{z}_2)
\end{aligned}
$$

4.2 FMCRT term

Averaging (4) over the BMCRT ensemble, i.e. over $p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) d\boldsymbol{y}_1 d\boldsymbol{y}_2$, one has

$$
\begin{aligned}
\langle C \rangle_B \;=\;& L_0(\boldsymbol{x}_{-1}, \boldsymbol{y}_0) \\
&+ K(\boldsymbol{y}_0 - \boldsymbol{x}_2) w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
&+ \int w_{2,1}(\boldsymbol{x}_0, \boldsymbol{y}_1) E(\boldsymbol{y}_0) L_0(\boldsymbol{y}_0, \boldsymbol{y}_1) p_B(\boldsymbol{y}_0, \boldsymbol{y}_1) d^2 \boldsymbol{y}_1 \\
&+ \int w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_1, \boldsymbol{y}_2) p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) d^2 \boldsymbol{y}_1 d^2 \boldsymbol{y}_2 \\
&+ \sum_{n \geq 3} K(\boldsymbol{y}_0 - \boldsymbol{x}_n) w_{3,0}(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2}) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
&+ \sum_{n \geq 2} w_{3,1}(\boldsymbol{y}_0, \boldsymbol{x}_n, \boldsymbol{x}_{n-1}) E(\boldsymbol{y}_0) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{x}_n}, \overrightarrow{\boldsymbol{x}_n \boldsymbol{y}_0}, \boldsymbol{x}_n) p_B(\boldsymbol{y}_0, \boldsymbol{x}_n) \\
&+ \sum_{n \geq 2} \int w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_n) E(\boldsymbol{y}_0, \boldsymbol{y}_1) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{x}_n}, \overrightarrow{\boldsymbol{x}_n \boldsymbol{y}_1}, \boldsymbol{x}_n) \\
&\qquad \times p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_n) d^2 \boldsymbol{y}_1
\end{aligned}
$$

Repeating the reasoning from Section **4.1**, we conclude that for small S the main contribution to $\langle \langle C \rangle_B^2 \rangle_F$ comes from squares of the kernel, so

$$
\begin{aligned}
\langle C \rangle_B^2 \;\approx\;& S^{-1} K(\boldsymbol{y}_0 - \boldsymbol{x}_2) w_{2,0}^2(\boldsymbol{y}_0, \boldsymbol{x}_1) f^2(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
&+ S^{-1} \sum_{n \geq 3} K(\boldsymbol{y}_0 - \boldsymbol{x}_n) w_{3,0}^2(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2}) f^2(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
&+ O(1)
\end{aligned}
$$

and averaging over the FMCRT ensemble, i.e. $p_F(\boldsymbol{x}_1, \boldsymbol{x}_2, ...) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \cdots$ one arrives at

$$
\begin{aligned}
\langle \langle C \rangle_B^2 \rangle_F \;\approx\;& S^{-1} \int w_0^2(\boldsymbol{y}_0, \boldsymbol{x}_1, \boldsymbol{x}_2) f^2(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \bar{p}_F(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{y}_0) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \\
&+ S^{-1} \int w_0^2(\boldsymbol{y}_0, \boldsymbol{x}_1) f^2(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) p_F(\boldsymbol{x}_1, \boldsymbol{y}_0) d\boldsymbol{x}_1 \\
&+ O(1)
\end{aligned}
$$

where $\bar{p}_F$ was defined in (44) and p_F in (41), so using (46)

$$
\begin{aligned}
\langle\langle C\rangle_B^2\rangle_F \;\approx\; & S^{-1}\int w_{2,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1)f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)\\
& \times L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right|s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& +S^{-1}\int w_{3,0}^2(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2\\
& +O(1) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (12)
\end{aligned}
$$

where $\boldsymbol{z}_0=\boldsymbol{y}_0$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ denote the vertices of the *joint* path counted from camera (regardless of which subpath they are from: camera or light). Notice these are the same terms with $w_{2,0}^2$ and $w_{3,0}^2$ as in (10).

4.3 BMCRT term

Introducing

$$
\begin{aligned}
\hat{F}_m \;\equiv\; & \int w_{2,m}(\boldsymbol{y}_0,\boldsymbol{z}_1)L_0(\boldsymbol{y}_0,\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{y}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0},\boldsymbol{y}_0)\\
& \times\left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0})\right|s(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \qquad\qquad\qquad (13)\\
F_m \;\equiv\; & \int w_{3,m}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2\\
\;=\; & \int G_m(\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0},\boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0})\right|s(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
G_m(\boldsymbol{z}_1) \;\equiv\; & \int w_{3,m}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)L(\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2},\overrightarrow{\boldsymbol{y}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\\
& \times\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right|s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2 \qquad\qquad\qquad (14)
\end{aligned}
$$

one can write (47) as (notice that by normalization conditions $w_{1,0}=1$):

$$
\begin{aligned}
\langle C\rangle_F \;=\; & \left(L_0(\boldsymbol{y}_{-1},\boldsymbol{y}_0)+\hat{F}_0+F_0\right)\\
& +w_{2,1}(\boldsymbol{y}_0,\boldsymbol{y}_1)E(\boldsymbol{y}_0)L_0(\boldsymbol{y}_0,\boldsymbol{y}_1)+E(\boldsymbol{y}_0)G_1(\boldsymbol{y}_1)\\
& +w_{3,2}(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{y}_2)E(\boldsymbol{y}_0,\boldsymbol{y}_1)L(\boldsymbol{y}_1,\boldsymbol{y}_2)
\end{aligned}
$$

Squaring it and averaging over BMCRT ensemble, i.e. over $p_B(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{y}_2)d\boldsymbol{y}_1 d\boldsymbol{y}_2$ gives

$$\langle\langle C\rangle_F^2\rangle_B = \left(L_0(\boldsymbol{z}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0\right)^2$$

$$+ \int w_{2,1}^2(\boldsymbol{z}_0, \boldsymbol{z}_1) E^2(\boldsymbol{z}_0) L_0^2(\boldsymbol{z}_0, \boldsymbol{z}_1) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ 2 \int w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) G_1(\boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ \int G_1^2(\boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ \int w_{3,2}^2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) L^2(\boldsymbol{z}_1, \boldsymbol{z}_2) p_B(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2$$

$$+ 2 \left(L_0(\boldsymbol{z}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0\right)\left(\hat{F}_1 + F_1 + F_2\right)$$

$$+ 2 \int w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) G_2(\boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ 2 \int G_1(\boldsymbol{z}_1) G_2(\boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1 \tag{15}$$

where $\boldsymbol{z}_0 = \boldsymbol{y}_0$, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ denote the vertices of the *joint* path counted from camera.

5   Optimal weights and variation of noise

Optimal weights are such that any change of them only increases the noise. In other words, these weights are an extremum of the noise functional. By definition, these are the weights which minimize the noise functional V, and thus the variation of it (caused by variation of weights) must vanish.

Since $\langle\langle C\rangle\rangle$ is independent of weight (provided that they sum to 1, see Appendix B), (5) implies that the variation of the pixel RMS is the sum of variations of the three averages, $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle\langle C\rangle_F^2\rangle_B$

$$\delta_m V = \frac{1}{N_B N_F}\delta_m\langle\langle C^2\rangle\rangle + \frac{\left(1 - N_B^{-1}\right)}{N_F}\delta_m\langle\langle C\rangle_B^2\rangle_F + \frac{\left(1 - N_F^{-1}\right)}{N_B}\delta_m\langle\langle C\rangle_F^2\rangle_B \tag{16}$$

where δ_m is variation with respect to the m-th weight. The optimal weights are those for which $\delta_m V$ for an arbitrary δw_m.

Because of the normalization condition (50) there are only 3 independent weights: one from the family $\{w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1), w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1)\}$ and two from the family $\{w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2), w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2), w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)\}$. We decided to choose

$$w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1), w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2), w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$$

When calculating the variation $\delta_1\langle\langle C^2\rangle\rangle$ against $w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$, one must remember that when $w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ is varied while $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ is kept, then (since the sum of weights is fixed), $w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ also varies:

$$w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \;\mapsto\; w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$$
$$w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \;\mapsto\; w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) - \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$$

and similarly, when we calculate the variation $\delta_2\langle\langle C^2\rangle\rangle$ against $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$. When $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ is varied the second weight of this family, i.e. $w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1)$, also changes by $-\delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$.

Now let us calculate variation of our three terms (10), (12), (15) with respect to these independent weights.

5.1 Cross term

The change of $\langle\langle C^2\rangle\rangle$ from (10) in response to the variation of $w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ (keeping $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ and $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ fixed) is

$$
\begin{aligned}
\delta_1\langle\langle C^2\rangle\rangle \;\approx\; &-2S^{-1}\int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) \\
&\times \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
&+2S^{-1}\int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E(\boldsymbol{z}_0) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1}, \boldsymbol{z}_1) \\
&\times \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
=\; &2S^{-1}\int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
&\times \Big(w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0) \\
&\qquad - w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0) \Big) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \quad (17)
\end{aligned}
$$

Here, as in the previous Section, $\boldsymbol{z}_0$, $\boldsymbol{z}_1$, $\boldsymbol{z}_1$ denote the vertices of the *joint* path counted from camera (regardless of which subpath they are from: camera or light).

Similarly, the change of $\langle\langle C^2\rangle\rangle$ in response to the variation of $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ (keeping $w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ and $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ fixed) is

$$\delta_2\langle\langle C^2\rangle\rangle \approx -2S^{-1}\int \delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)$$
$$\times \rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2$$
$$+2S^{-1}\int \delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}$$
$$\times \rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2$$
$$= 2S^{-1}\int \delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times \left(w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)} \right.$$
$$\left. -w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0) \right)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \qquad (18)$$

At last, variation $\delta w_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)$ results in

$$\delta_0\langle\langle C^2\rangle\rangle \approx 2S^{-1}\int \delta w_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)w_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$\times f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right|s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1 \quad (19)$$

5.2 FMCRT term

The change of $\langle\langle C\rangle_B^2\rangle_F$ from (12) in response to the variation of $w_{3,1}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$ (keeping $w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$ and $w_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)$ fixed so $\delta w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = -\delta w_{3,1}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$) is

$$\delta_1\langle\langle C\rangle_B^2\rangle_F \approx -2S^{-1}\int \delta w_{3,1}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \qquad (20)$$

Similarly, variation $\delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$ results in

$$\delta_2\langle\langle C\rangle_B^2\rangle_F \approx -2S^{-1}\int \delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \qquad (21)$$

Variation $\delta w_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)$ results in

$$\delta_0\langle\langle C^2\rangle\rangle \approx 2S^{-1}\int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)f^2(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0}, \boldsymbol{z}_0)$$
$$\times L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right|s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1 \tag{22}$$

5.3 BMCRT term

Variation $\delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ results in (recall that since $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ is kept fixed, $w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ also varies by $-\delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$)

$$
\begin{aligned}
\delta_1\langle\langle C\rangle_F^2\rangle_B &= 2\left(L_0(\boldsymbol{y}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0\right)F[\delta w_{3,1}] \\
&\quad -2\left(L_0(\boldsymbol{y}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0\right)F[\delta w_{3,1}] \\
&\quad -2\left(\hat{F}_1 + F_1 + F_2\right)F[\delta w_{3,1}] \\
&\quad +2\int w_1(\boldsymbol{z}_0, \boldsymbol{z}_1)L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)G[\delta w_{3,1}](\boldsymbol{z}_1)E^2(\boldsymbol{z}_0)p_B(\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
&\quad +2\int(G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1))G[\delta w_{3,1}]E^2(\boldsymbol{z}_0)p_B(\boldsymbol{z}_1)d\boldsymbol{z}_1 \\
&= -2F_\Sigma F[\delta w_{3,1}] \\
&\quad +2\int(w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1)L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1)) \\
&\quad \times(G[\delta w_{3,1}])(\boldsymbol{z}_1)E^2(\boldsymbol{z}_0)p_B(\boldsymbol{z}_1)d\boldsymbol{z}_1
\end{aligned}
$$

where
$$F_\Sigma \equiv \hat{F}_1 + F_1 + F_2 \tag{23}$$

and

$$
\begin{aligned}
F[h] &\equiv \int h(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)\rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
&= \int(G[h])(\boldsymbol{z}_1)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \overrightarrow{\boldsymbol{z}_0\boldsymbol{y}_{-1}}, \boldsymbol{z}_0)\left|(\boldsymbol{n}(\boldsymbol{z}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0})\right|s(\boldsymbol{z}_0, \boldsymbol{z}_1)d\boldsymbol{z}_1 \\
(G[h])(\boldsymbol{z}_1) &\equiv \int h(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2 L(\boldsymbol{z}_1, \boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0}, \boldsymbol{z}_1) \\
&\quad \times\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right|s(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2 \tag{24}
\end{aligned}
$$

generalize (14) by integral operators acting on an arbitrary function of 3 arguments $h(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$.

After some trivial algebra this $\delta_1\langle\langle C\rangle_F^2\rangle_B$ can be written as

$$
\begin{aligned}
\delta_1 \langle \langle C \rangle_F^2 \rangle_B \;=\; & -2 F_\Sigma \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) L(\boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{y}_{-1}}, \boldsymbol{z}_0) f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) \\
& \times (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0})(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}) s(\boldsymbol{z}_0, \boldsymbol{z}_1) s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& +2 \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E(\boldsymbol{z}_0, \boldsymbol{z}_1) \\
& \times (w_1(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1))\, L(\boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{z}_0 \boldsymbol{y}_{-1}}, \boldsymbol{z}_0) \\
& \times (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0})(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}) s(\boldsymbol{z}_0, \boldsymbol{z}_1) s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
\;=\; & 2 \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times \Bigg(E(\boldsymbol{z}_0, \boldsymbol{z}_1)\, (w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1)) \\
& \qquad - F_\Sigma \Bigg) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \qquad\qquad (25)
\end{aligned}
$$

Similarly, variation $\delta w_2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ results in

$$
\begin{aligned}
\delta_2 \langle \langle C \rangle_F^2 \rangle_B \;=\; & -2 \left(L_0(\boldsymbol{y}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0 \right) F[\delta w_{3,2}] \\
& +2 \left(L_0(\boldsymbol{y}_{-1}, \boldsymbol{z}_0) + \hat{F}_0 + F_0 \right) F[\delta w_{3,2}] \\
& -2 F_\Sigma F[\delta w_{3,2}] \\
& +2 \int (w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1)) \\
& \quad \times (G[\delta w_{3,2}])\, (\boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) d\boldsymbol{z}_1 \\
& +2 \int \delta w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \\
& \quad \times L^2(\boldsymbol{z}_1, \boldsymbol{z}_2) p_B(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
\;=\; & 2 \int \delta w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times \Bigg(E(\boldsymbol{z}_0, \boldsymbol{z}_1)\, (w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1)) \\
& \qquad + w_2(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) E(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) L(\boldsymbol{z}_1, \boldsymbol{z}_2) - F_\Sigma \Bigg) d\boldsymbol{z}_1 d\boldsymbol{z}_2 (26)
\end{aligned}
$$

Variation $\delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ results in

$$
\begin{aligned}
\delta_0 \langle\langle C\rangle_F^2\rangle_B \;=\; & +2F_\Sigma \hat{F}[\delta w_{2,0}] \\
& -2\int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) E^2(\boldsymbol{z}_0) p_B(\boldsymbol{z}_1) \\
& \times (G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) + w_1(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1))\, d\boldsymbol{z}_1 \\
=\; & 2\int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) \\
& \times (F_\Sigma - (G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) + w_1(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1))\, E(\boldsymbol{z}_0)) \\
& \times f(\overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)\, \big| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}) \big|\, s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1 \qquad (27)
\end{aligned}
$$

where

$$
\hat{F}[h] \equiv \int h(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)\, \big| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}) \big|\, s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1
\tag{28}
$$

generalizes $\hat{F}_1$ from (13) by the integral operator acting on an arbitrary function $h(\boldsymbol{z}_0, \boldsymbol{z}_1)$ of two arguments.

5.4 Summing all up: variation of the total noise

Substituting variations of $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle\langle C\rangle_F^2\rangle_B$ from (17), (20) and (25) into (16) and neglecting $N_F^{-1} \ll 1$, we have

$$
\begin{aligned}
\frac{\delta_1 V}{2} \;\approx\; & \frac{1}{N_B n_F} \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times \Bigg(w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0) \\
& \qquad - w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) \Bigg) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& -\frac{1 - N_B^{-1}}{n_F} \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& +\frac{1}{N_B} \int \delta w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \rho(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) \\
& \times (E(\boldsymbol{z}_0)\, (w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) + G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1)) - F_\Sigma)\, d\boldsymbol{z}_1 d\boldsymbol{z}_2
\end{aligned}
$$

where $n_F \equiv S N_F$. This variation vanishes for an arbitrary $\delta w_{3,1}$ if and only if

$$w_{3,1}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\boldsymbol{z}_1)E(\boldsymbol{z}_0)-N_B w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)$$
$$+n_F w_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)E(\boldsymbol{z}_0)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$= n_F\left(F_\Sigma - E(\boldsymbol{z}_0)\left(G_1(\boldsymbol{z}_1)+G_2(\boldsymbol{z}_1)\right)\right) \tag{29}$$

Similarly, substituting variations of $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle\langle C\rangle_F^2\rangle_B$ from (18), (21) and (26) into (16) and neglecting $N_F^{-1}\ll 1$, we have

$$\frac{\delta_2 V}{2}\approx\frac{1}{N_B n_F}\int\delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times\left(w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}\right.$$
$$\left.-w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)\right)dz_1 dz_2$$
$$-\frac{1-N_B^{-1}}{n_F}\int\delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)w_0(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)dz_1 dz_2$$
$$+\frac{1}{N_B}\int\delta w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\rho(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)$$
$$\times\left(E(\boldsymbol{z}_0)\left(w_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)+G_1(\boldsymbol{z}_1)\right)\right.$$
$$\left.+w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)L(\boldsymbol{z}_1,\boldsymbol{z}_2)-F_\Sigma\right)dz_1 dz_2$$

which vanishes for an arbitrary δw_2 if and only if

$$w_{3,2}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{z}_0,\boldsymbol{z}_1)c(\boldsymbol{z}_1,\boldsymbol{z}_2)-N_B w_{3,0}(\boldsymbol{z}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0)$$
$$+n_F w_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)E(\boldsymbol{z}_0)L_0(\boldsymbol{z}_0,\boldsymbol{z}_1)$$
$$= n_F\left(F_\Sigma - E(\boldsymbol{z}_0)G_1(\boldsymbol{z}_1)\right) \tag{30}$$

where
$$c(\boldsymbol{z}_1,\boldsymbol{z}_2)\equiv\frac{b(\boldsymbol{z}_1,\boldsymbol{z}_2)}{L(\boldsymbol{z}_1,\boldsymbol{z}_2)}+n_F L(\boldsymbol{z}_1,\boldsymbol{z}_2)$$

and the integrals G and F are given by (14).

Similarly, substituting variations of $\langle\langle C^2\rangle\rangle$, $\langle\langle C\rangle_B^2\rangle_F$ and $\langle\langle C\rangle_F^2\rangle_B$ from (19), (22) and (27) into (16) and neglecting $N_F^{-1}\ll 1$, we have

$$\frac{\delta_0 V}{2} = \frac{1}{N_B n_F} \int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

$$\times f^2(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) \left| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}) \right| s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ \frac{1 - N_B^{-1}}{n_F} \int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

$$\times f^2(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) \left| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}) \right| s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1$$

$$+ \frac{1}{N_B} \int \delta w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

$$\times \left(F_\Sigma - (G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) + w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)) E(\boldsymbol{z}_0) \right)$$

$$\times f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) \left| (\boldsymbol{n}(\boldsymbol{z}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}) \right| s(\boldsymbol{z}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1$$

which vanishes for an arbitrary $\delta w_0(\boldsymbol{z}_0, \boldsymbol{z}_1)$ if and only if

$$\left(N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0) + n_F L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) E(\boldsymbol{z}_0) \right) w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$$

$$= n_F \left((G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) + L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)) E(\boldsymbol{z}_0) - F_\Sigma \right)$$

which yields

$$w_{2,0}(\boldsymbol{y}_0, \boldsymbol{z}_1) = W_{2,0}(\boldsymbol{y}_0, \boldsymbol{z}_1)$$

$$+ n_F \frac{(G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1)) E(\boldsymbol{y}_0) - F_\Sigma}{N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{y}_0}, \boldsymbol{y}_0) + n_F L_0(\boldsymbol{y}_0, \boldsymbol{z}_1) E(\boldsymbol{y}_0)} \quad (31)$$

$$W_{2,0}(\boldsymbol{y}_0, \boldsymbol{z}_1) \equiv \frac{n_F L_0(\boldsymbol{y}_0, \boldsymbol{z}_1) E(\boldsymbol{y}_0)}{N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{y}_0}, \boldsymbol{y}_0) + n_F L_0(\boldsymbol{y}_0, \boldsymbol{z}_1) E(\boldsymbol{y}_0)} \quad (32)$$

$$W_{2,1}(\boldsymbol{y}_0, \boldsymbol{z}_1) \equiv \frac{N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{y}_0}, \boldsymbol{y}_0)}{N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{y}_0}, \boldsymbol{y}_0) + n_F L_0(\boldsymbol{y}_0, \boldsymbol{z}_1) E(\boldsymbol{y}_0)}$$

Substituting $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ from (31) into the (29) and (30) and recalling that $w_{3,0}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) = 1 - w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) - w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ one has

$$\left(f(\overrightarrow{\overline{z_2 z_1}}, \overrightarrow{\overline{z_1 z_0}}, z_1) E(z_0) + N_B f(\overrightarrow{\overline{z_1 z_0}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) \right) w_{3,1}(z_0, z_1, z_2)$$
$$+ N_B w_{3,2}(z_0, z_1, z_2) f(\overrightarrow{\overline{z_1 z_0}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0)$$
$$= N_B f(\overrightarrow{\overline{z_1 z_0}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) - n_F E(z_0) L_0(z_0, z_1) W_{2,1}(z_0, z_1)$$
$$+ n_F \left(F_\Sigma - E(z_0) \left(G_1(z_1) + G_2(z_1) \right) \right) W_{2,1}(z_0, z_1)$$

$$N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) w_{3,1}(z_0, z_1, z_2)$$
$$+ w_{3,2}(z_0, z_1, z_2) \left(E(z_0, z_1) c(z_1, z_2) + N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) \right)$$
$$= N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) - n_F E(z_0) L_0(z_0, z_1) W_{2,1}(z_0, z_1)$$
$$+ n_F W_{2,1}(z_0, z_1) F_\Sigma - n_F E(z_0) W_{2,1}(z_0, z_1) G_1(z_1)$$
$$+ n_F E(z_0) W_{2,0}(z_0, z_1) G_2(z_1)$$

Denoting

$$A_{11}(z_0, z_1, z_2) \equiv f(\overrightarrow{\overline{z_2 z_1}}, \overrightarrow{\overline{z_1 z_0}}, z_1) E(z_0) + N_B f(\overrightarrow{\overline{z_1 z_0}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0)$$
$$A_{12}(z_0, z_1) \equiv N_B f(\overrightarrow{\overline{z_1 z_0}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0)$$
$$A_{22}(z_0, z_1, z_2) \equiv E(z_0, z_1) c(z_1, z_2) + N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0)$$
$$\Phi(z_0, z_1) \equiv N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) - n_F W_{2,1}(z_0, z_1) E(z_0) L_0(z_0, z_1)$$
$$= N_B f(\overrightarrow{\overline{z_0 z_1}}, \overrightarrow{\overline{y_{-1} z_0}}, z_0) W_{2,1}(z_0, z_1)$$

it takes the form

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix} \begin{pmatrix} w_{3,1}(z_0, z_1, z_2) \\ w_{3,2}(z_0, z_1, z_2) \end{pmatrix} = \left(\begin{array}{c} \Phi(z_1) + n_F W_{2,1}(z_0, z_1) F_\Sigma \\ \\ -n_F E(z_0) W_{2,1}(z_0, z_1) G_1(z_1) \end{array} \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$-n_F E(z_0) G_2(z_1) \begin{pmatrix} W_{2,1}(z_0, z_1) \\ -W_{2,0}(z_0, z_1) \end{pmatrix}$$

so

$$
\begin{aligned}
w_{3,1} &= W_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + n_F W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{A_{22} - A_{12}}{A_{11}A_{22} - A_{12}^2} F_\Sigma \\
&\quad - n_F E(\boldsymbol{z}_0) W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{A_{22} - A_{12}}{A_{11}A_{22} - A_{12}^2} G_1(\boldsymbol{z}_1) \\
&\quad - n_F E(\boldsymbol{z}_0) \frac{A_{22} W_{2,1} + A_{12} W_{2,0}}{A_{11}A_{22} - A_{12}^2} G_2(\boldsymbol{z}_1) \quad (33) \\
w_{3,2} &= W_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) + n_F W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{A_{11} - A_{12}}{A_{11}A_{22} - A_{12}^2} F_\Sigma \\
&\quad - n_F E(\boldsymbol{z}_0) W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) \frac{A_{11} - A_{12}}{A_{11}A_{22} - A_{12}^2} G_1(\boldsymbol{z}_1) \\
&\quad + n_F E(\boldsymbol{z}_0) \frac{A_{11} W_{2,0} + A_{12} W_{2,1}}{A_{11}A_{22} - A_{12}^2} G_2(\boldsymbol{z}_1) \quad (34)
\end{aligned}
$$

where

$$
\begin{aligned}
W_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &\equiv \frac{A_{22} - A_{12}}{A_{11}A_{22} - A_{12}^2} \Phi \\
&= \frac{1}{1 + \frac{n_F L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)}} \\
&\quad \times \frac{1}{1 + \frac{f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0)}{N_B f(\overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)} + \frac{f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0)}{E(\boldsymbol{z}_0, \boldsymbol{z}_1) c(\boldsymbol{z}_1, \boldsymbol{z}_2)}} \quad (35) \\
W_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) &\equiv \frac{A_{11} - A_{12}}{A_{11}A_{22} - A_{12}^2} \Phi \\
&= \frac{1}{1 + \frac{n_F L_0(\boldsymbol{z}_0, \boldsymbol{z}_1)}{N_B \tilde{f}(\overrightarrow{\boldsymbol{z}_0 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)}} \\
&= \frac{\frac{f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0)}{E(\boldsymbol{z}_0, \boldsymbol{z}_1) c(\boldsymbol{z}_1, \boldsymbol{z}_2)}}{1 + \frac{f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0)}{N_B f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \overrightarrow{\boldsymbol{y}_{-1} \boldsymbol{z}_0}, \boldsymbol{z}_0)} + \frac{f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{z}_0}, \boldsymbol{z}_1) E(\boldsymbol{z}_0)}{E(\boldsymbol{z}_0, \boldsymbol{z}_1) c(\boldsymbol{z}_1, \boldsymbol{z}_2)}} \quad (36)
\end{aligned}
$$

Here

$$
\tilde{f}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}) \equiv \frac{f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z})}{\int f(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{z}) \left| (\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z})) \right| d^2 \boldsymbol{v}} \quad (37)
$$

is the normalized BDF used in our BMCRT without ray killing (all rays are scattered).

6 Solution of the integral equations in weights

Its integral terms G_1, G_2 and F are averages that reduce the number of arguments and this allows to find the solution in a close form.

6.1 The idea of solution

The idea is very simple. Let us consider a simplified equation of the same type:

$$w(\boldsymbol{z}_1, \boldsymbol{z}_2) = W(\boldsymbol{z}_1, \boldsymbol{z}_2) + a(\boldsymbol{z}_1, \boldsymbol{z}_2)G(\boldsymbol{z}_1)$$
$$G(\boldsymbol{z}_1) \equiv \int b(\boldsymbol{z}_1, \boldsymbol{z}_2)w(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2$$

where w is the sought-for function while a and b and W are known. Multiplying this equation by $b(\boldsymbol{z}_1, \boldsymbol{z}_2)$ and integrating over $\boldsymbol{z}_2$ we have

$$G(\boldsymbol{z}_1) = \int b(\boldsymbol{z}_1, \boldsymbol{z}_2)W(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2 + G(\boldsymbol{z}_1)\int a(\boldsymbol{z}_1, \boldsymbol{z}_2)b(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2$$

because g depends only on $\boldsymbol{z}$ and thus can be moved outside the integral. This is the principal feature allowing solution in a closed form. From this equation we immediately find that

$$G(\boldsymbol{z}_1) = \frac{\int b(\boldsymbol{z}_1, \boldsymbol{z}_2)W(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2}{1 - \int a(\boldsymbol{z}_1, \boldsymbol{z}_2)b(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2}$$

and then

$$w(\boldsymbol{z}_1, \boldsymbol{z}_2) = W(\boldsymbol{z}_1, \boldsymbol{z}_2) + \frac{a(\boldsymbol{z}_1, \boldsymbol{z}_2)\int b(\boldsymbol{z}_1, \boldsymbol{z}_2)W(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2}{1 - \int a(\boldsymbol{z}_1, \boldsymbol{z}_2)b(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_2}$$

6.2 Application of the method to the real system

Now let us apply this idea to the system in weights (33)–(34) which is slightly more complex because the solution is a 3-vector and it also contains the integrals of $G_m(\boldsymbol{z}_1)$i.e. F_m defined by 14. This makes the solution more tedious though the driving idea is the same. According to this idea we substitute $w_{3,1}$ and $w_{3,2}$ from (33)–(34) into the definition (14) of G_m which gives

$$
\begin{aligned}
G_1 \;=\; & \left(\mathcal{A}_{22}(\boldsymbol{z}_1) - \mathcal{A}_{12}(\boldsymbol{z}_1)\right)\left(\Phi(\boldsymbol{z}_0,\boldsymbol{z}_1) + n_F W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)F_\Sigma\right) \\
& -n_F E(\boldsymbol{z}_0)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(\mathcal{A}_{22}-\mathcal{A}_{12}\right)G_1(\boldsymbol{z}_1) \\
& -n_F E(\boldsymbol{z}_0)\left(\mathcal{A}_{22}W_{2,1}+\mathcal{A}_{12}W_{2,0}\right)G_2(\boldsymbol{z}_1) \\
G_2 \;=\; & \left(\mathcal{A}_{11}(\boldsymbol{z}_1) - \mathcal{A}_{12}(\boldsymbol{z}_1)\right)\left(\Phi(\boldsymbol{z}_0,\boldsymbol{z}_1) + n_F W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)F_\Sigma\right) \\
& -n_F E(\boldsymbol{z}_0)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(\mathcal{A}_{11}-\mathcal{A}_{12}\right)G_1(\boldsymbol{z}_1) \\
& +n_F E(\boldsymbol{z}_0)\left(\mathcal{A}_{11}W_{2,0}+\mathcal{A}_{12}W_{2,1}\right)G_2(\boldsymbol{z}_1)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{A}_{mn}(\boldsymbol{z}_1) \;\equiv\; & \int \frac{A_{mn}}{A_{11}A_{22}-A_{12}^2}L(\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_0},\boldsymbol{z}_1) \\
& \times(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2
\end{aligned}
$$

We therefore obtain a 2x2 linear system in G_1, G_2:

$$
\begin{aligned}
& \begin{pmatrix} B_{11}(\boldsymbol{z}_1) & B_{11}(\boldsymbol{z}_1) \\ B_{21}(\boldsymbol{z}_1) & B_{22}(\boldsymbol{z}_1) \end{pmatrix} \begin{pmatrix} G_1(\boldsymbol{z}_1) \\ G_2(\boldsymbol{z}_1) \end{pmatrix} \\
& = \left(\Phi(\boldsymbol{z}_0,\boldsymbol{z}_1) + n_F F_\Sigma W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) \begin{pmatrix} H_1(\boldsymbol{z}_1) \\ H_2(\boldsymbol{z}_1) \end{pmatrix}
\end{aligned} \tag{38}
$$

where

$$
\begin{aligned}
B_{11}(\boldsymbol{z}_1) \;\equiv\; & 1 + n_F E(\boldsymbol{z}_0)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(\mathcal{A}_{22}(\boldsymbol{z}_1)-\mathcal{A}_{12}(\boldsymbol{z}_1)\right) \\
B_{12}(\boldsymbol{z}_1) \;\equiv\; & n_F E(\boldsymbol{z}_0)\left(\mathcal{A}_{22}(\boldsymbol{z}_1)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)+\mathcal{A}_{12}(\boldsymbol{z}_1)W_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) \\
B_{21}(\boldsymbol{z}_1) \;\equiv\; & n_F E(\boldsymbol{z}_0)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)\left(\mathcal{A}_{11}(\boldsymbol{z}_1)-\mathcal{A}_{12}(\boldsymbol{z}_1)\right) \\
B_{22}(\boldsymbol{z}_1) \;\equiv\; & 1 - n_F E(\boldsymbol{z}_0)\left(\mathcal{A}_{12}(\boldsymbol{z}_1)W_{2,1}(\boldsymbol{z}_0,\boldsymbol{z}_1)+\mathcal{A}_{11}(\boldsymbol{z}_1)W_{2,0}(\boldsymbol{z}_0,\boldsymbol{z}_1)\right) \\
H_1(\boldsymbol{z}_1) \;\equiv\; & \mathcal{A}_{22}(\boldsymbol{z}_1)-\mathcal{A}_{12}(\boldsymbol{z}_1) \\
H_2(\boldsymbol{z}_1) \;\equiv\; & \mathcal{A}_{11}(\boldsymbol{z}_1)-\mathcal{A}_{12}(\boldsymbol{z}_1)
\end{aligned}
$$

Therefore

$$
\begin{aligned}
G_m(\boldsymbol{z}_1) \;=\; & \left(\Phi(\boldsymbol{z}_0,\boldsymbol{z}_1) + n_F F_\Sigma W_1(\boldsymbol{z}_0,\boldsymbol{z}_1)\right)\mathfrak{H}_m(\boldsymbol{z}_1) \\
\;=\; & \left(N_B f(\overrightarrow{\boldsymbol{z}_0\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{z}_0},\boldsymbol{z}_0) + n_F F_\Sigma\right)W_1(\boldsymbol{z}_0,\boldsymbol{z}_1)\mathfrak{H}_m(\boldsymbol{z}_1)
\end{aligned} \tag{39}
$$

where

$$
\begin{pmatrix} \mathfrak{H}_1(\boldsymbol{z}_1) \\ \mathfrak{H}_2(\boldsymbol{z}_1) \end{pmatrix} \equiv \begin{pmatrix} B_{11}(\boldsymbol{z}_1) & B_{12}(\boldsymbol{z}_1) \\ B_{21}(\boldsymbol{z}_1) & B_{22}(\boldsymbol{z}_1) \end{pmatrix}^{-1} \begin{pmatrix} H_1(\boldsymbol{z}_1) \\ H_2(\boldsymbol{z}_1) \end{pmatrix}
$$

Then, according to definitions (14), (13) of F_m and $\hat{F}_1$ one has

$$
\begin{aligned}
F_\Sigma &= F_1 + F_2 + \hat{F}_1 \\
&= \int \left(G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) \right) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{n}(\boldsymbol{z}_0) \cdot \boldsymbol{v}) V d^2\boldsymbol{v} \\
&\quad + \int w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1(\boldsymbol{v})) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v}
\end{aligned}
$$

where $\boldsymbol{u}_0 \equiv \overrightarrow{\boldsymbol{y}_{-1}\boldsymbol{y}_0}$ and $\boldsymbol{v} \equiv \overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0}$; naturally $d^2\boldsymbol{v} = s(\boldsymbol{y}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1$.

Substituting $w_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) = 1 - w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ with $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ from (31) becomes

$$
\begin{aligned}
F_\Sigma &= \int W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} \\
&\quad + \int W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) \left(G_1(\boldsymbol{z}_1) + G_2(\boldsymbol{z}_1) \right) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} \\
&\quad + F_\Sigma \frac{1}{E(\boldsymbol{z}_0)} \int W_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v}
\end{aligned}
$$

Substituting here $G_m(\boldsymbol{z}_1)$ from (39) we arrive at

$$
\begin{aligned}
F_\Sigma &= \int W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} \\
&\quad + N_B \int W_{2,1}^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \left(\mathfrak{H}_1(\boldsymbol{z}_1) + \mathfrak{H}_2(\boldsymbol{z}_1) \right) f^2(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} \\
&\quad + n_F F \int W_{2,1}^2(\boldsymbol{z}_0, \boldsymbol{z}_1) \left(\mathfrak{H}_1(\boldsymbol{z}_1) + \mathfrak{H}_2(\boldsymbol{z}_1) \right) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} \\
&\quad + F_\Sigma \frac{1}{E(\boldsymbol{z}_0)} \int W_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v}
\end{aligned}
$$

which implies

$$
F_\Sigma(\boldsymbol{z}_0) = \frac{\int W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) L_0(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} + N_B D_2(\boldsymbol{z}_0)}{\frac{1}{E(\boldsymbol{z}_0)} \int W_{2,1}(\boldsymbol{z}_0, \boldsymbol{z}_1) f(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0)) d^2\boldsymbol{v} - n_F D_1(\boldsymbol{z}_0)}
$$

where

$$D_m(\boldsymbol{z}_0) \equiv \int W_{2,1}^2(\boldsymbol{z}_0, \boldsymbol{z}_1)\,(\mathfrak{H}_1(\boldsymbol{z}_1) + \mathfrak{H}_2(\boldsymbol{z}_1))\,f^m(\boldsymbol{v}, \boldsymbol{u}_0, \boldsymbol{z}_0)(\boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{z}_0))d^2\boldsymbol{v}$$

Here all the functions are known (they depend only on BRDF and illumination, but not on the weights).

We then substitute this F_Σ into (39) and $G_1(\boldsymbol{z})$ and $G_2(\boldsymbol{z})$ *become known.* Then they together with F_Σ are in their turn substituted into (33)–(34) and we obtain $w_{3,1}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$ and $w_{3,2}(\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2)$.

At last we substitute the above F_Σ and $G_1(\boldsymbol{z})$ and $G_2(\boldsymbol{z})$ into (31) and obtain $w_{2,0}(\boldsymbol{z}_0, \boldsymbol{z}_1)$ which completes the solution.

7 Conclusion

We investigated how the "partial MIS" which mixes just two strategies for BDPM [10] and extends to the case of mixing three strategies. It happens that this case has several qualitative differences from mixing only two strategies. Meanwhile further increase of the number of strategies past three already brings no qualitatively changes.

We derive the system of integral equations which couples weights from the two different families. This system extends the single integral equation in the single independent weight from [10]. It has a special type which admits solution in close form, i.e. as an analytic formula which includes some integrals of known functions that must be calculated numerically. Qualitatively these integrals are similar to those which arise while mixing two strategies, and therefore also can be calculated during ray tracing.

The above system of integral equations replaces the algebraic system of the famous Veach heuristic [9, 4] for the optimal weights in MCRT. In the bidirectional ray tracing the noise functional is different from that in MCRT and therefore the system for weights differs from the Veach heuristic. For example, while the latter depends only on BDFs at the path vertices, our formulae also include illumination (or luminance) of that points. Also our system of equations is not local, i.e. the weights for a given path depend on the scene properties not only at the vertices of that path but outside of it also.

A Appendix. MCRT densities

The densities p_F and p_B can be calculated from the simple recurrence relations (similar to that in [4]).

Below we shall sometimes use the spatial and sometimes angular probability densities keeping in mind the obvious relation between differentials (8)

A.1 BMCRT density

In our variant of BMCRT direction $\boldsymbol{u}$ of the next segment $[\boldsymbol{y}_m, \boldsymbol{y}_{m+1}]$ of the camera ray is chosen with probability density proportional to the normalized BDF to *exclude ray killing,* i.e. with probability density $\tilde{f}(\boldsymbol{u}, \overrightarrow{\boldsymbol{y}_{m-1}\boldsymbol{y}_m}, \boldsymbol{y}_m)\left|(\boldsymbol{n}(\boldsymbol{y}_m) \cdot \boldsymbol{u}\right|$, where $\tilde{f}$ is BDF normalized for BMCRT (37). Therefore, the probability density for the path obeys the recurrence relation

$$
\begin{aligned}
p_B(\boldsymbol{y}_0, ..., \boldsymbol{y}_{m+1}) = \ & p_B(\boldsymbol{y}_0, ..., \boldsymbol{y}_m) \\
& \times \tilde{f}(\overrightarrow{\boldsymbol{y}_{m+1}\boldsymbol{y}_m}, \overrightarrow{\boldsymbol{y}_{m-1}\boldsymbol{y}_m}, \boldsymbol{y}_m)\left|(\boldsymbol{n}(\boldsymbol{y}_m) \cdot \overrightarrow{\boldsymbol{y}_{m+1}\boldsymbol{y}_m})\right| s(\boldsymbol{y}_m, \boldsymbol{y}_{m+1}),
\end{aligned}
$$

where we dropped the fixed point $\boldsymbol{y}_{-1}$ which is the same for all pixels = the camera origin. In other words,

$$
\begin{aligned}
p_B(\boldsymbol{y}_0, \boldsymbol{y}_1) = \ & \tilde{f}(\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0})\right| s(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
= \ & \frac{1}{E(\boldsymbol{y}_0)} f(\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1})\right| s(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
p_B(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) = \ & p_B(\boldsymbol{y}_0, \boldsymbol{y}_1) \times \tilde{f}(\overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1}, \boldsymbol{y}_1)\left|(\overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1} \cdot \boldsymbol{n}(\boldsymbol{y}_1))\right| s(\boldsymbol{y}_1, \boldsymbol{y}_2) \\
= \ & \frac{1}{E(\boldsymbol{y}_0, \boldsymbol{y}_1)} f(\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1})\right| f(\overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1}, \boldsymbol{y}_1) \\
& \times \left|(\overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1} \cdot \boldsymbol{n}(\boldsymbol{y}_1))\right| s(\boldsymbol{y}_0, \boldsymbol{y}_1) s(\boldsymbol{y}_1, \boldsymbol{y}_2) \\
= \ & \tilde{f}(\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)\tilde{f}(\overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1}, \boldsymbol{y}_1) \\
& \times \frac{\left|(\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0})\right|\left|(\boldsymbol{n}(\boldsymbol{y}_1) \cdot \overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1})\right|}{|\boldsymbol{y}_0 - \boldsymbol{y}_1|^2} \\
& \times \frac{\left|(\boldsymbol{n}(\boldsymbol{y}_1) \cdot \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_1})\right|\left|(\boldsymbol{n}(\boldsymbol{y}_2) \cdot \overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_2})\right|}{|\boldsymbol{y}_1 - \boldsymbol{y}_2|^2}
\end{aligned}
\tag{40}
$$

A.2 FMCRT density

In our variant of FMCRT with its Russian roulette, direction $\boldsymbol{v}$ of the next segment $[\boldsymbol{x}_m, \boldsymbol{x}_{m+1}]$ of light ray is chosen with probability density proportional to BDF (not normalized to provide *ray killing*) $f(\overrightarrow{\boldsymbol{x}_{m-1}\boldsymbol{x}_m}, \boldsymbol{v}, \boldsymbol{x}_m)\left|(\boldsymbol{n}(\boldsymbol{x}_m) \cdot \boldsymbol{v})\right|$. Therefore, the probability density for the path obeys the recurrence relation

$$
\begin{aligned}
p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_{m+1}) = \ & p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_m) \\
& \times f(\overrightarrow{\boldsymbol{x}_{m-1}\boldsymbol{x}_m}, \overrightarrow{\boldsymbol{x}_m\boldsymbol{x}_{m+1}}, \boldsymbol{x}_m)\left|(\boldsymbol{n}(\boldsymbol{x}_m) \cdot \overrightarrow{\boldsymbol{x}_m\boldsymbol{x}_{m+1}})\right| s(\boldsymbol{x}_m, \boldsymbol{x}_{m+1})
\end{aligned}
$$

where we dropped the fixed point $\boldsymbol{x}_0$, i.e. the light source position.

The first term of this recurrence $p_F(\boldsymbol{x}_1, \boldsymbol{x}_2)$ can be also calculated from another considerations. It is the density for the light (sub)path $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \boldsymbol{x}_2$ in which $\boldsymbol{x}_1$ is illuminated *directly*. Meanwhile, by definition $L(\boldsymbol{x}_2, \boldsymbol{x}_1)(\boldsymbol{n}(\boldsymbol{x}_1) \cdot \overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1})$ is the *intensity* of emission from $\boldsymbol{x}_1$ towards $\boldsymbol{x}_2$ which in FMCRT equals the (angular) density of segments $[\boldsymbol{x}_1, \boldsymbol{x}_2]$ rays. So,

$$
\begin{aligned}
p_F(\boldsymbol{x}_1, \boldsymbol{x}_2) &= L_0(\boldsymbol{x}_2, \boldsymbol{x}_1)(\boldsymbol{n}(\boldsymbol{x}_1) \cdot \overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1})s(\boldsymbol{x}_1, \boldsymbol{x}_2) \\
&= L_0(\boldsymbol{x}_2, \boldsymbol{x}_1)(\boldsymbol{n}(\boldsymbol{x}_2) \cdot \overrightarrow{\boldsymbol{x}_2\boldsymbol{x}_1})s(\boldsymbol{x}_2, \boldsymbol{x}_1) \qquad (41)
\end{aligned}
$$

where $L_0(\boldsymbol{x}_2, \boldsymbol{x}_1)$ is the *direct* luminance of the point $\boldsymbol{x}_1$ towards $\boldsymbol{x}_2$ and $s(\boldsymbol{x}_1, \boldsymbol{x}_2)$ transforms it into the spatial density.

The above recurrence can also be written as

$$
\begin{aligned}
p_F(\boldsymbol{X}^{(m)}, \boldsymbol{y}) &= p_F(\boldsymbol{X}^{(m)}) \\
&\times f(\overrightarrow{\boldsymbol{x}_{m-1}\boldsymbol{x}_m}, \overrightarrow{\boldsymbol{x}_m\boldsymbol{y}}, \boldsymbol{x}_m)\left|(\boldsymbol{n}(\boldsymbol{x}_m) \cdot \overrightarrow{\boldsymbol{x}_m\boldsymbol{y}})\right| s(\boldsymbol{x}_m, \boldsymbol{y}) \qquad (42)
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{X}^{(0)} &\equiv null \\
\boldsymbol{X}^{(m)} &\equiv (\boldsymbol{x}_1, ..., \boldsymbol{x}_m), \quad m = 1, 2, ... \qquad (43)
\end{aligned}
$$

is light sub-path and $\boldsymbol{y}$ is the *next* hit point.

Besides $p_F(\boldsymbol{x}_1, ...)$ which accounts for the whole light path, we shall also need the density $\bar{p}_F$ of its "end" (regardless of the initial part):

$$
\begin{aligned}
\bar{p}_F(\boldsymbol{z}_1, \boldsymbol{z}_2) &\equiv \sum_{m \geq 0} \int p_F(\boldsymbol{X}^{(m)}, \boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{X}^{(m)} \\
\bar{p}_F(\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3) &\equiv \sum_{m \geq 0} \int p_F(\boldsymbol{X}^{(m)}, \boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3)d\boldsymbol{X}^{(m)} \qquad (44)
\end{aligned}
$$

We used notations $\boldsymbol{z}_1, ...$ for the tail vertices because $\boldsymbol{x}_1$ is reserved for the first hit point etc.

By construction $\bar{p}_F(\boldsymbol{z}_1, \boldsymbol{z}_2)d\boldsymbol{z}_1d\boldsymbol{z}_2$ is the fraction of the light rays that hit the surface element $d\boldsymbol{z}_1$ near $\boldsymbol{z}_1$ and then hit the surface element $d\boldsymbol{z}_2$ near $\boldsymbol{z}_2$. On the other hand $L(\boldsymbol{z}_2, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2})\right| d\boldsymbol{z}_1$ is the *angular* density of light rays emitted from $\boldsymbol{z}_1$ towards $\boldsymbol{z}_2$. Therefore

$$
\bar{p}_F(\boldsymbol{z}_1, \boldsymbol{z}_2) = L(\boldsymbol{z}_2, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2})\right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) \qquad (45)
$$

Notice that since $\bar{p}_F$ includes *all* rays that hit near $\boldsymbol{z}_1$ and go towards $\boldsymbol{z}_2$, L is the *full* luminance (direct+indirect).

Also, applying (42) to (44) we have

$$
\begin{aligned}
\bar{p}_F(\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3) \;=\;& L(\boldsymbol{z}_2, \boldsymbol{z}_1)\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2})\right| \\
&\times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{z}_2}, \overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_3}, \boldsymbol{z}_2)s(\boldsymbol{z}_1, \boldsymbol{z}_2)s(\boldsymbol{z}_2, \boldsymbol{z}_3) \qquad (46)
\end{aligned}
$$

B Appendix. Normalization condition (sum of weights)

Let us calculate the limiting luminance of the pixel $\langle\langle C\rangle_F\rangle_B$ as a functional on weights. Averaging the contribution C given by (4) over the FMCRT ensemble we have

$$
\begin{aligned}
\langle C\rangle_F \;=\;& L_0(\boldsymbol{y}_{-1}, \boldsymbol{y}_0)w_{1,0}(\boldsymbol{y}_0) \\
&+E(\boldsymbol{x}_0)L_0(\boldsymbol{x}_0, \boldsymbol{x}_1)w_{2,1}(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
&+\int K(\boldsymbol{y}_0-\boldsymbol{x}_2)w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1)f(\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)p_F(\boldsymbol{x}_1, \boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2 \\
&+E(\boldsymbol{y}_0, \boldsymbol{y}_1)L_0(\boldsymbol{y}_1, \boldsymbol{y}_2)w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) \\
&+\sum_{n\geq 3}\int K(\boldsymbol{y}_0-\boldsymbol{x}_n)w_{3,0}(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2}) \\
&\qquad\times f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}}, \boldsymbol{y}_0)p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)d\boldsymbol{x}_1\cdots d\boldsymbol{x}_n \\
&+\sum_{n\geq 2}\int K(\boldsymbol{y}_1-\boldsymbol{x}_n)w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_{n-1})E(\boldsymbol{y}_0) \\
&\qquad\times f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0}, \boldsymbol{y}_1)p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)d\boldsymbol{x}_1\cdots d\boldsymbol{x}_n \\
&+\sum_{n\geq 2}\int K(\boldsymbol{y}_2-\boldsymbol{x}_n)w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2)E(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
&\qquad\times f(\overrightarrow{\boldsymbol{x}_{n-1}\boldsymbol{x}_n}, \overrightarrow{\boldsymbol{y}_2\boldsymbol{y}_1}, \boldsymbol{y}_2)p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)d\boldsymbol{x}_1\cdots d\boldsymbol{x}_n
\end{aligned}
$$

where

$$
p_F(\boldsymbol{x}_1, ..., \boldsymbol{x}_n)
$$

is the density (not normalized, i.e. with account for absorption!) at the light (sub)path $\boldsymbol{x}_0 \to \boldsymbol{x}_1 \to \cdots \to \boldsymbol{x}_n$ where we drop the *fixed* light source position $\boldsymbol{x}_0$.

Neglecting the difference between the kernel and the delta-function, we arrive at

$$
\begin{aligned}
\langle C \rangle_F \approx\ & L_0(\boldsymbol{y}_{-1}, \boldsymbol{y}_0) w_{1,0}(\boldsymbol{y}_0) \\
& + E(\boldsymbol{y}_0) L_0(\boldsymbol{y}_0, \boldsymbol{y}_1) w_{2,1}(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
& + \int w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) p_F(\boldsymbol{x}_1, \boldsymbol{y}_0) d\boldsymbol{x}_1 \\
& + E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_1, \boldsymbol{y}_2) w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) \\
& + \sum_{n \geq 3} \int w_{3,0}(\boldsymbol{y}_0, \boldsymbol{x}_{n-1}, \boldsymbol{x}_{n-2}) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
& \quad \times p_F(\boldsymbol{X}^{(n-3)}, \boldsymbol{x}_{n-2}, \boldsymbol{x}_{n-1}, \boldsymbol{y}_0) d\boldsymbol{X}^{(n-3)} d\boldsymbol{x}_{n-2} d\boldsymbol{x}_{n-1} \\
& + \sum_{n \geq 2} \int w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{x}_{n-1}) E(\boldsymbol{y}_0) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_1 \boldsymbol{y}_0}, \boldsymbol{y}_1) \\
& \quad \times p_F(\boldsymbol{X}^{(n-2)}, \boldsymbol{x}_{n-1}, \boldsymbol{y}_1) d\boldsymbol{X}^{(n-2)} d\boldsymbol{x}_{n-1} \\
& + \sum_{n \geq 2} \int w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) f(\overrightarrow{\boldsymbol{x}_{n-1} \boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2 \boldsymbol{y}_1}, \boldsymbol{y}_2) \\
& \quad \times p_F(\boldsymbol{X}^{(n-2)}, \boldsymbol{x}_{n-1}, \boldsymbol{y}_2) d\boldsymbol{X}^{(n-2)} d\boldsymbol{x}_{n-1}
\end{aligned}
$$

(where $\boldsymbol{X}^{(m)}$ is the starting part of the light path, see (43)), or

$$
\begin{aligned}
\langle C \rangle_F \approx\ & L_0(\boldsymbol{y}_{-1}, \boldsymbol{y}_0) w_{1,0}(\boldsymbol{y}_0) \\
& + E(\boldsymbol{y}_0) L_0(\boldsymbol{y}_0, \boldsymbol{y}_1) w_{2,1}(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
& + \int w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) p_F(\boldsymbol{x}_1, \boldsymbol{y}_0) d\boldsymbol{x}_1 \\
& + E(\boldsymbol{y}_0, \boldsymbol{y}_1) L(\boldsymbol{y}_1, \boldsymbol{y}_2) w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) \\
& + \int w_{3,0}(\boldsymbol{y}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
& \qquad \times \left(\sum_{m \geq 0} \int p_F(\boldsymbol{X}^{(m)}, \boldsymbol{z}_2, \boldsymbol{z}_1, \boldsymbol{y}_0) d\boldsymbol{X}^{(m)} \right) d\boldsymbol{z}_1 d\boldsymbol{z}_2 \\
& + \int w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{z}_2) E(\boldsymbol{y}_0) f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_1 \boldsymbol{y}_0}, \boldsymbol{y}_1) \\
& \qquad \times \left(\sum_{m \geq 0} \int p_F(\boldsymbol{X}^{(m)}, \boldsymbol{z}_2, \boldsymbol{y}_1) d\boldsymbol{X}^{(m)} \right) d\boldsymbol{z}_2 \\
& + w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) \int f(\overrightarrow{\boldsymbol{z}_3 \boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2 \boldsymbol{y}_1}, \boldsymbol{y}_2) \\
& \qquad \times \left(\sum_{m \geq 0} \int p_F(\boldsymbol{X}^{(m)}, \boldsymbol{z}_3, \boldsymbol{y}_2) d\boldsymbol{X}^{(m)} \right) d\boldsymbol{z}_3
\end{aligned}
$$

where $\boldsymbol{z}_{-1}$, $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ denote the vertices of the *joint* path counted from camera (regardless of which subpath they are from: camera or light). Obviously, always $\boldsymbol{z}_{-1} = \boldsymbol{y}_{-1}$ is the camera origin and $\boldsymbol{z}_0 = \boldsymbol{y}_0$ is the first hit which is fixed for the given pixel.

Using (42) and (44), (45), (46) and (41) we have

$$
\begin{aligned}
\langle C \rangle_F \approx\ & L_0(\boldsymbol{y}_{-1}, \boldsymbol{y}_0) w_{1,0}(\boldsymbol{y}_0) \\
& + w_{2,1}(\boldsymbol{y}_0, \boldsymbol{y}_1) E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
& + \int w_{2,0}(\boldsymbol{y}_0, \boldsymbol{x}_1) L_0(\boldsymbol{y}_0, \boldsymbol{x}_1) f(\overrightarrow{\boldsymbol{x}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \\
& \qquad \times \left| (\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{y}_0 \boldsymbol{x}_1}) \right| s(\boldsymbol{y}_0, \boldsymbol{x}_1) d\boldsymbol{x}_1 \\
& + w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_1, \boldsymbol{y}_2) \\
& + \int \Bigg(\int w_{3,0}(\boldsymbol{y}_0, \boldsymbol{z}_1, \boldsymbol{z}_2) L(\boldsymbol{z}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}, \overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \boldsymbol{z}_1) \\
& \qquad \times \left| (\boldsymbol{n}(\boldsymbol{z}_1) \cdot \overrightarrow{\boldsymbol{z}_2 \boldsymbol{z}_1}) \right| s(\boldsymbol{z}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \Bigg) \\
& \qquad \times f(\overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}, \overrightarrow{\boldsymbol{y}_0 \boldsymbol{y}_{-1}}, \boldsymbol{y}_0) \left| (\boldsymbol{n}(\boldsymbol{y}_0) \cdot \overrightarrow{\boldsymbol{z}_1 \boldsymbol{y}_0}) \right| s(\boldsymbol{y}_0, \boldsymbol{z}_1) d\boldsymbol{z}_1 \\
& + E(\boldsymbol{y}_0) \int w_{3,1}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{z}_2) L(\boldsymbol{y}_1, \boldsymbol{z}_2) f(\overrightarrow{\boldsymbol{z}_2 \boldsymbol{y}_1}, \overrightarrow{\boldsymbol{y}_1 \boldsymbol{y}_0}, \boldsymbol{y}_1) \\
& \qquad \times \left| (\boldsymbol{n}(\boldsymbol{y}_1) \cdot \overrightarrow{\boldsymbol{z}_2 \boldsymbol{y}_1}) \right| s(\boldsymbol{y}_1, \boldsymbol{z}_2) d\boldsymbol{z}_2 \\
& + w_2(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) \\
& \qquad \times \int f(\overrightarrow{\boldsymbol{z}_3 \boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2 \boldsymbol{y}_1}, \boldsymbol{y}_2) L(\boldsymbol{y}_2, \boldsymbol{z}_3) \left| (\boldsymbol{n}(\boldsymbol{y}_2) \cdot \overrightarrow{\boldsymbol{z}_3 \boldsymbol{y}_2}) \right| s(\boldsymbol{y}_2, \boldsymbol{z}_3) d\boldsymbol{z}_3 \quad (47)
\end{aligned}
$$

The last integral is nothing but the diffuse luminance at $\boldsymbol{x}_2$ towards $\boldsymbol{x}_1$:

$$
\int f(\overrightarrow{\boldsymbol{z}_3 \boldsymbol{y}_2}, \overrightarrow{\boldsymbol{y}_2 \boldsymbol{y}_1}, \boldsymbol{y}_2) L(\boldsymbol{y}_2, \boldsymbol{z}_3) \left| (\boldsymbol{n}(\boldsymbol{y}_2) \cdot \overrightarrow{\boldsymbol{z}_3 \boldsymbol{y}_2}) \right| s(\boldsymbol{y}_2, \boldsymbol{z}_3) d\boldsymbol{z}_3 = L_d(\boldsymbol{y}_1, \boldsymbol{y}_2)
$$

so its sum with $w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) L_0(\boldsymbol{y}_1, \boldsymbol{y}_2)$ gives $w_{3,2}(\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2) E(\boldsymbol{y}_0, \boldsymbol{y}_1) L(\boldsymbol{y}_1, \boldsymbol{y}_2)$.

Now let us average over the BMCRT ensemble, i.e. over the points $\boldsymbol{y}_1, \boldsymbol{y}_2$ ($\boldsymbol{y}_0$ is fixed for given pixel). Then (47) becomes

$$
\begin{aligned}
\langle\langle C\rangle_F\rangle_B \;\approx\; & L_0(\boldsymbol{y}_{-1},\boldsymbol{y}_0)w_{1,0}(\boldsymbol{y}_0)\\
& + \int w_{2,1}(\boldsymbol{y}_0,\boldsymbol{x}_1)L_0(\boldsymbol{y}_0,\boldsymbol{x}_1)E(\boldsymbol{y}_0)p_B(\boldsymbol{y}_0,\boldsymbol{x}_1)d\boldsymbol{x}_1\\
& + \int w_{2,0}(\boldsymbol{y}_0,\boldsymbol{x}_1)L_0(\boldsymbol{y}_0,\boldsymbol{x}_1)f(\overrightarrow{\boldsymbol{x}_1\boldsymbol{y}_0},\overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}},\boldsymbol{y}_0)\\
& \quad \times \left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{y}_0\boldsymbol{x}_1})\right|s(\boldsymbol{y}_0,\boldsymbol{x}_1)d\boldsymbol{x}_1\\
& + \int w_{3,2}(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{y}_2)L(\boldsymbol{y}_1,\boldsymbol{y}_2)E(\boldsymbol{y}_0,\boldsymbol{y}_1)p_B(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{y}_2)d\boldsymbol{y}_1 d\boldsymbol{y}_2\\
& + \int\left(\int w_{3,0}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)L(\boldsymbol{z}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0},\boldsymbol{z}_1)\right.\\
& \quad \times\left.\left|(\boldsymbol{n}(\boldsymbol{z}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1})\right|s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2\right)\\
& \quad \times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0},\overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}},\boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0})\right|s(\boldsymbol{z}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& + E(\boldsymbol{y}_0)\int\left(\int w_{3,1}(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{z}_2)L(\boldsymbol{y}_1,\boldsymbol{z}_2)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{y}_1},\overrightarrow{\boldsymbol{y}_1\boldsymbol{y}_0},\boldsymbol{y}_1)\right.\\
& \quad \times\left.\left|(\boldsymbol{n}(\boldsymbol{y}_1)\cdot\overrightarrow{\boldsymbol{z}_2\boldsymbol{y}_1})\right|s(\boldsymbol{y}_1,\boldsymbol{z}_2)d\boldsymbol{z}_2\right)p_B(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1
\end{aligned}
$$

where $p_B(\boldsymbol{y}_0,\boldsymbol{y}_1)$ and $p_B(\boldsymbol{y}_0,\boldsymbol{y}_1,\boldsymbol{y}_2)$ are the density at the camera (subpath) $\boldsymbol{y}_{-1}\to\boldsymbol{y}_0\to\boldsymbol{y}_1$ or $\boldsymbol{y}_{-1}\to\boldsymbol{y}_0\to\boldsymbol{y}_1\to\boldsymbol{y}_2$, respectively. So,

$$
\begin{aligned}
\langle\langle C\rangle_F\rangle_B \;\approx\; & w_{1,0}(\boldsymbol{y}_0)L_0(\boldsymbol{y}_{-1},\boldsymbol{y}_0)\\
& + \int\left(w_{2,0}(\boldsymbol{y}_0,\boldsymbol{z}_1)+w_{2,1}(\boldsymbol{y}_0,\boldsymbol{z}_1)\right)L_0(\boldsymbol{y}_0,\boldsymbol{z}_1)\\
& \quad \times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0},\overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}},\boldsymbol{y}_0)\left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{y}_0\boldsymbol{z}_1})\right|s(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1\\
& + \int\left(w_{3,0}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)+w_{3,1}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)+w_{3,2}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)\right)L(\boldsymbol{z}_1,\boldsymbol{z}_2)\\
& \quad \times f(\overrightarrow{\boldsymbol{z}_1\boldsymbol{y}_0},\overrightarrow{\boldsymbol{y}_0\boldsymbol{y}_{-1}},\boldsymbol{y}_0)f(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1},\overrightarrow{\boldsymbol{y}_0\boldsymbol{z}_1},\boldsymbol{z}_1)\\
& \quad \times \left|(\boldsymbol{n}(\boldsymbol{y}_0)\cdot\overrightarrow{\boldsymbol{y}_0\boldsymbol{z}_1})\right|\left|(\overrightarrow{\boldsymbol{z}_2\boldsymbol{z}_1}\cdot\boldsymbol{n}(\boldsymbol{z}_1))\right|s(\boldsymbol{y}_0,\boldsymbol{z}_1)s(\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2
\end{aligned}
$$

which can be written as

$$\langle\langle C\rangle_F\rangle_B \approx w_{1,0}(\boldsymbol{y}_0)L_0(\boldsymbol{y}_{-1},\boldsymbol{y}_0)$$
$$+\int (w_{2,0}(\boldsymbol{y}_0,\boldsymbol{z}_1) + w_{2,1}(\boldsymbol{y}_0,\boldsymbol{z}_1))\, L_0(\boldsymbol{y}_0,\boldsymbol{z}_1)E(\boldsymbol{y}_0)p_B(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1$$
$$+\int (w_{3,0}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) + w_{3,1}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) + w_{3,2}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2))$$
$$\times L(\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{y}_0,\boldsymbol{z}_1)p_B(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \tag{48}$$

The "pure" BDD=2 corresponds to weights

$$w_{1,0}(\boldsymbol{y}_0) = 1$$
$$w_{2,0}(\boldsymbol{y}_0,\boldsymbol{z}_1) = 0$$
$$w_{2,1}(\boldsymbol{y}_0,\boldsymbol{z}_1) = 1$$
$$w_{3,0}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = 0$$
$$w_{3,1}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = 0$$
$$w_{3,2}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = 1$$

so for BDD=2 the expectation of the pixel luminance is

$$\langle\langle C\rangle_F\rangle_B \approx L_0(\boldsymbol{y}_{-1},\boldsymbol{y}_0)$$
$$+\int L_0(\boldsymbol{y}_0,\boldsymbol{z}_1)E(\boldsymbol{y}_0)p_B(\boldsymbol{y}_0,\boldsymbol{z}_1)d\boldsymbol{z}_1$$
$$+\int L(\boldsymbol{z}_1,\boldsymbol{z}_2)E(\boldsymbol{y}_0,\boldsymbol{z}_1)p_B(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2)d\boldsymbol{z}_1 d\boldsymbol{z}_2 \tag{49}$$

Meanwhile, the strategy BDD=2 gives a correct result, i.e. $\langle\langle C\rangle_F\rangle_B$ converges to the exact pixel luminance as $R \to 0$. Comparing (48) and (49) we see that the 3-point mixed strategy gives correct result if

$$w_{2,0}(\boldsymbol{y}_0) = 1$$
$$w_{2,1}(\boldsymbol{y}_0,\boldsymbol{z}_1) + w_{2,0}(\boldsymbol{y}_0,\boldsymbol{z}_1) = 1$$
$$w_{3,0}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) + w_{3,1}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) + w_{3,2}(\boldsymbol{y}_0,\boldsymbol{z}_1,\boldsymbol{z}_2) = 1 \tag{50}$$

Notice that formally this is a sufficient, not a necessary condition.

C   Appendix. Average of the product of kernels

Indeed, we have the following terms to be neglected:

Terms with the same $\boldsymbol{x}$ but different $\boldsymbol{y}$

$$K(\boldsymbol{y}_0 - \boldsymbol{x}_n)K(\boldsymbol{y}_1 - \boldsymbol{x}_n), \quad n \geq 2$$
$$K(\boldsymbol{y}_0 - \boldsymbol{x}_n)K(\boldsymbol{y}_2 - \boldsymbol{x}_n), \quad n \geq 2$$
$$K(\boldsymbol{y}_1 - \boldsymbol{x}_n)K(\boldsymbol{y}_2 - \boldsymbol{x}_n), \quad n \geq 2$$

Terms with different $\boldsymbol{x}$ but same $\boldsymbol{y}$; for all of them $n \neq m, \quad m \geq 2$

$$K(\boldsymbol{y}_0 - \boldsymbol{x}_n)K(\boldsymbol{y}_0 - \boldsymbol{x}_m), \quad n \geq 3$$
$$K(\boldsymbol{y}_1 - \boldsymbol{x}_n)K(\boldsymbol{y}_1 - \boldsymbol{x}_m), \quad n \geq 2$$
$$K(\boldsymbol{y}_2 - \boldsymbol{x}_n)K(\boldsymbol{y}_2 - \boldsymbol{x}_m), \quad n \geq 2$$

Terms with different $\boldsymbol{x}$ and different $\boldsymbol{y}$; for all of them $n \neq m$,

$$K(\boldsymbol{y}_0 - \boldsymbol{x}_n)K(\boldsymbol{y}_1 - \boldsymbol{x}_m), \quad n, m \geq 3$$
$$K(\boldsymbol{y}_0 - \boldsymbol{x}_n)K(\boldsymbol{y}_2 - \boldsymbol{x}_m), \quad n, m \geq 3$$
$$K(\boldsymbol{y}_1 - \boldsymbol{x}_n)K(\boldsymbol{y}_2 - \boldsymbol{x}_m), \quad n, m \geq 2$$

Notice that in the integral these products of kernels are multiplied by a smooth functions which in all cases we shall denote $a(\boldsymbol{X}; \boldsymbol{Y})$ where $\boldsymbol{X} \equiv (\boldsymbol{x}_1, ..., \boldsymbol{x}_n, ...,)$ is light path and $\boldsymbol{Y} \equiv (\boldsymbol{y}_0, ..., \boldsymbol{y}_i, ...,)$ is camera path.

Obviously for $K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_i - \boldsymbol{x}_m)$ assuming $m > n$ the average over the FMCRT ensemble is

$$\int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_i - \boldsymbol{x}_m)a(\boldsymbol{X}; \boldsymbol{Y})d\boldsymbol{X}$$
$$\approx \int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_i - \boldsymbol{x}_m)d\boldsymbol{x}_n d\boldsymbol{x}_m$$
$$\times \int a(\boldsymbol{x}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n+1}, ..., \boldsymbol{x}_{m-1}, \boldsymbol{y}_i, ...; \boldsymbol{Y})d\boldsymbol{X}$$
$$= \int a(\boldsymbol{x}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n+1}, ..., \boldsymbol{x}_{m-1}, \boldsymbol{y}_i, ...; \boldsymbol{Y})d\boldsymbol{X}$$

i.e. asymptotically independent of S. So this integral is $O(1)$.

Similarly, for $K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_m)$ assuming $m > n$ the average over FMCRT ensemble is

$$\int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_m)a(\boldsymbol{X};\boldsymbol{Y})d\boldsymbol{X}$$

$$\approx \int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_m)d\boldsymbol{x}_n d\boldsymbol{x}_m$$

$$\times \int a(\boldsymbol{x}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n+1}, ..., \boldsymbol{x}_{m-1}, \boldsymbol{y}_j, ...; \boldsymbol{Y})d\boldsymbol{X}$$

$$= \int a(\boldsymbol{x}_1, ..., \boldsymbol{y}_i, \boldsymbol{x}_{n+1}, ..., \boldsymbol{x}_{m-1}, \boldsymbol{y}_j, ...; \boldsymbol{Y})d\boldsymbol{X}$$

i.e. again asymptotically independent of S. So this integral is $O(1)$.

At last, for $K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_n)$ the average over the FMCRT ensemble is

$$\int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_n)a(\boldsymbol{X};\boldsymbol{Y})d\boldsymbol{X}$$

$$\approx \int K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_n)d\boldsymbol{x}_n \times \int a(\boldsymbol{x}_1, ..., \boldsymbol{x}_{n-1}, \boldsymbol{y}_i, ...; \boldsymbol{Y})d\boldsymbol{X}$$

Let us calculate

$$\int K(\boldsymbol{y}_i - \boldsymbol{z})K(\boldsymbol{y}_j - \boldsymbol{z})d\boldsymbol{z}$$

Obviously it equals S^{-2} (squared amplitude of the kernel) times *area of intersection* of two circles of radii R with centres at $\boldsymbol{y}_i$ and at $\boldsymbol{y}_j$ where $S = \pi R^2$. The latter was calculated in [17] and using this result,

$$\int K(\boldsymbol{y}_i - \boldsymbol{z})K(\boldsymbol{y}_j - \boldsymbol{z})d\boldsymbol{z} = S^{-2}2R^2\left(\arccos\frac{l}{2R} - \frac{l}{2R}\sqrt{1 - \frac{l}{2R}^2}\right) \quad (51)$$

where $l \equiv \left|\boldsymbol{y}_i - \boldsymbol{y}_j\right|$.

The average over the BMCRT ensemble then makes it

$$\int \left(\int K(\boldsymbol{y}_i - \boldsymbol{z})K(\boldsymbol{y}_j - \boldsymbol{z})d\boldsymbol{z}d\boldsymbol{y}_j \times \int a(\boldsymbol{x}_1, ..., \boldsymbol{x}_{n-1}, \boldsymbol{x}_i, \boldsymbol{x}_{n+1}, ...; \boldsymbol{Y})d\boldsymbol{X}\right)d\boldsymbol{y}_i$$

The integration over $\boldsymbol{y}_j$ is over the whole plane, and using (51) we can calculate it:

$$\int K(\boldsymbol{y}_i - \boldsymbol{z})K(\boldsymbol{y}_j - \boldsymbol{z})d\boldsymbol{z}d\boldsymbol{y}_j$$

$$= S^{-2}2R^2 \int \left(\arccos \frac{l}{2R} - \frac{l}{2R}\sqrt{1 - \frac{l}{2R}^2} \right) d\boldsymbol{y}_j$$

$$= \frac{2}{\pi S} \times 2\pi \int_0^{2R} \left(\arccos \frac{l}{2R} - \frac{l}{2R}\sqrt{1 - \frac{l}{2R}^2} \right) l\,dl$$

$$= \frac{16}{\pi} \int_0^1 \left(\arccos \frac{l}{2R} - \frac{l}{2R}\sqrt{1 - \frac{l}{2R}^2} \right) \frac{l}{2R}d\frac{l}{2R}$$

$$= \frac{16}{\pi} \int_0^1 \left(\arccos t - t\sqrt{1 - t^2} \right) t\,dt$$

$$= 1$$

(for fixed $\boldsymbol{y}_i$) so

$$\langle\langle K(\boldsymbol{y}_i - \boldsymbol{x}_n)K(\boldsymbol{y}_j - \boldsymbol{x}_n)a(\boldsymbol{X}, \boldsymbol{Y}\rangle_F\rangle_B$$

$$\approx \int a(\boldsymbol{x}_1, ..., \boldsymbol{x}_{n-1}, \boldsymbol{y}_i, \boldsymbol{x}_{n+1}, ...; \boldsymbol{Y})d\boldsymbol{X}d\boldsymbol{y}_i$$

$$= O(1)$$

Therefore, *all* the cross terms containing products of *different* kernel functions are $O(1)$. Meanwhile those containing a *square* of kernel function which equals $S^{-1}K$ are naturally $O(S^{-1})$, i.e. dominate.

References

[1] M. Šik and J. Křivánek, "Survey of Markov chain Monte Carlo methods in light transport simulation," *IEEE Trans. Visual. Comput. Graph.*, vol. 26, no. 4, pp. 1821–1840, 2018.

[2] M. Pharr and G. Humphreys, *Physically Based Rendering, Second Edition: From Theory To Implementation*, 2nd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2010.

[3] H. W. Jensen and P. Christensen, "High quality rendering using ray tracing and photon mapping," in *ACM SIGGRAPH 2007*

Courses. New York, NY, USA: ACM, 2007. [Online]. Available: http://doi.acm.org/10.1145/1281500.1281593

[4] J. Vorba, "Bidirectional photon mapping," in *Proceedings of CESCG 2011: The 15th Central European Seminar on Computer Graphics.* Prague: Charles University, 2011, pp. 25–32. [Online]. Available: https://cgg.mff.cuni.cz/ jaroslav/papers/2011-bdpm/vorba2011-bdpm.pdf

[5] I. Georgiev, J. Křivánek, T. Davidovič, and P. Slusallek, "Light transport simulation with vertex connection and merging," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 192:1–192:10, 2012. [Online]. Available: http://doi.acm.org/10.1145/2366145.2366211

[6] T. Hachisuka, J. Pantaleoni, and H. W. Jensen, "A path space extension for robust light transport simulation," *ACM Trans. Graph.*, vol. 31, no. 6, pp. 191:1–191:10, 2012.

[7] S. V. Ershov, D. D. Zhdanov, and A. G. Voloboy, "Estimation of noise in calculation of scattering medium luminance by MCRT," *Mathematica Montisnigri*, vol. 45, pp. 60–73, 2019. [Online]. Available: http://doi.acm.org/10.20948/mathmontis-2019-45-5

[8] M. Sbert, V. Havran, and L. Szirmay-Kalos, "Multiple importance sampling revisited: breaking the bounds," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, pp. 1–15, 2018.

[9] E. Veach, "Robust Monte-Carlo methods for light transport simulation," 1997. [Online]. Available: http://graphics.stanford.edu/papers/veach_thesis/thesis.pdf

[10] S. Ershov and A. Voloboy, "Calculation of MIS weights for bidirectional path tracing with photon maps in presence of direct illumination," *Mathematica Montisnigri*, vol. 48, pp. 86–102, 2020. [Online]. Available: http://doi.acm.org/10.20948/mathmontis-2020-48-8

[11] N. Dodik, "Implementing probabilistic connections for bidirectional path tracing in the Mitsuba Renderer," Favoritenstrasse 9-11/186, A-1040 Vienna, Austria, Sep. 2017. [Online]. Available: https://www.cg.tuwien.ac.at/research/publications/2017/dodik-2017-pcbpt/

[12] A. Keller, P. Grittmann, J. Vorba, I. Georgiev, M. Šik, E. d'Eon, P. Gautron, P. Vévoda, and I. Kondapaneni, "Advances in Monte Carlo

rendering: The legacy of Jaroslav Křivánek," in *ACM SIGGRAPH 2020 Courses*, New York, NY, USA, 2020, pp. 1–366. [Online]. Available: https://doi.org/10.1145/3388769.3407458

[13] I. Kondapaneni, P. Vévoda, P. Grittmann, T. Skřivan, P. Slusallek, and J. Křivánek, "Optimal multiple importance sampling," *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2019)*, vol. 38, no. 4, pp. 37:1–37:14, 2019.

[14] S. V. Ershov, E. D. Birukov, and A. G. Voloboy, "Efficient calculation of the optimal MIS weights in bi-directional ray tracing with photon maps," *Keldysh Institute Preprints*, no. 107, 2020. [Online]. Available: https://doi.org/10.20948/prepr-2020-107

[15] V. Havran, R. Herzog, and H. Seidel, "Fast final gathering via reverse photon mapping," *Computer Graphics Forum*, vol. 24, no. 3, pp. 323–333, 2005.

[16] A. Zhdanov, D. Zhdanov, and V. Galaktionov, "Realistic image synthesis with hybrid photon maps," *Proceedings of SPIE, Optoelectronic Imaging and Multimedia Technology VII*, vol. 10550, no. 115500G, 2020.

[17] S. Ershov, A. Voloboy, B. Barladian, V. Sokolov, I. Potemin, and D. Zhdanov, "Light scattering in automotive paints: continuous medium approach vs correlations between particles," *Proceedings of SPIE, Illumination Optics VI*, vol. 11874, no. 1187408, 2021. [Online]. Available: http://doi.acm.org/10.1117/12.2597166