

Integral formulations of volumetric transmittance

ILYAN GEORGIEV*, Autodesk, United Kingdom

ZACKARY MISSO*, Dartmouth College, USA

TOSHIYA HACHISUKA, The University of Tokyo, Japan

DEREK NOWROUZEZAHRAI, McGill University, Canada

JAROSLAV KŘIVÁNEK, Charles University and Chaos Czech a. s., Czech Republic

WOJCIECH JAROSZ, Dartmouth College, USA

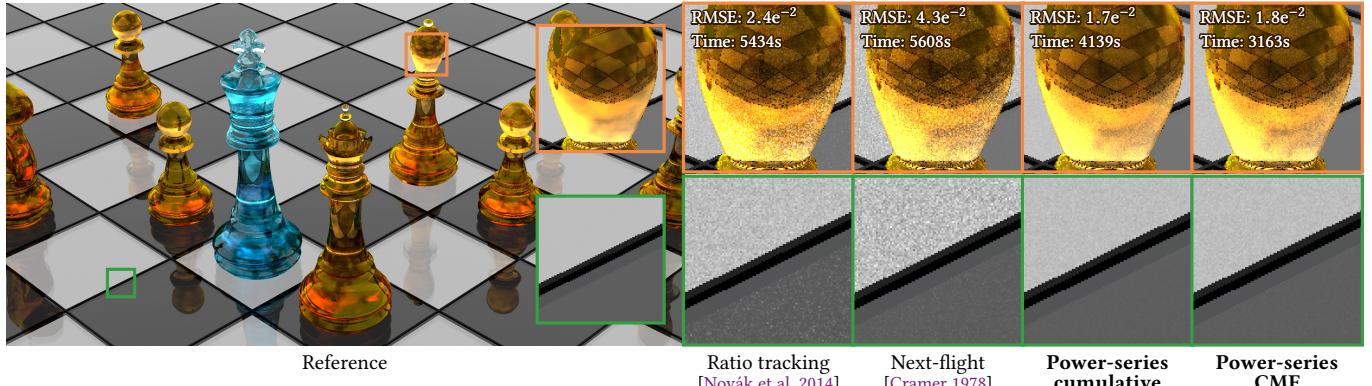


Fig. 1. Equal-lookup comparisons of various estimators in a scene with both homogeneous (chess board) and heterogeneous (chess pieces) media. Our proposed estimators (in bold) trade more extinction lookups for lower-variance estimates of transmittance. RMSE and render times are for the entire image.

Computing the light attenuation between two given points is an essential yet expensive task in volumetric light transport simulation. Existing unbiased transmittance estimators are all based on “null-scattering” random walks enabled by augmenting the media with fictitious matter. This formulation prevents the use of traditional Monte Carlo estimator variance analysis, thus the efficiency of such methods is understood from a mostly empirical perspective. In this paper, we present several novel integral formulations of volumetric transmittance in which existing estimators arise as direct Monte Carlo estimators. Breaking from physical intuition, we show that the null-scattering concept is not strictly required for unbiased transmittance estimation, but is a form of control variates for effectively reducing variance. Our formulations bring new insight into the problem and the efficiency of existing estimators. They also provide a framework for devising new types of transmittance estimators with distinct and complementary performance tradeoffs, as well as a clear recipe for applying sample stratification.

CCS Concepts: • Computing methodologies → Rendering; Ray tracing.

Additional Key Words and Phrases: participating media, transmittance, null collision, null scattering, stochastic sampling, Monte Carlo integration

ACM Reference Format:

Iliyan Georgiev, Zackary Misso, Toshiya Hachisuka, Derek Nowrouzezahrai, Jaroslav Křivánek, and Wojciech Jarosz. 2019. Integral formulations of volumetric transmittance. *ACM Trans. Graph.* 38, 6, Article 154 (November 2019), 17 pages. <https://doi.org/10.1145/3355089.3356559>

1 INTRODUCTION

The accurate reproduction of lighting in participating media, such as clouds, fog, or smoke, is important in realistic image synthesis as well as other fields, ranging from medical imaging to fire safety and building design. A vast body of research has been devoted to efficient volumetric light transport simulation over the past two decades [Cerezo et al. 2005; Novák et al. 2018]. State-of-the-art Monte Carlo methods sample a large number of random light trajectories between sensors and emitters to approximate the light transport in the scene. These include (bidirectional) path tracing, many-light methods, and Metropolis light transport; for a comprehensive overview we refer the reader to Novák et al. [2018].

Common to all these approaches is the task of computing the transmittance, i.e. fractional visibility, between any two points in the scene. The transmittance is traditionally formulated as an exponential of the medium optical thickness, i.e. integrated extinction, along the line segment connecting the points. Computing the transmittance is trivial when the optical thickness integral takes an analytic form, e.g. in homogeneous media. Heterogeneous media, however,

*Authors with equal contribution.

Authors' addresses: Iliyan Georgiev, iliyan.georgiev@autodesk.com, Autodesk, London, United Kingdom; Zackary Misso, zackary.t.misso@dartmouth.edu, Dartmouth College, Hanover, NH, USA; Toshiya Hachisuka, thachisuka@siggraph.org, The University of Tokyo, Tokyo, Japan; Derek Nowrouzezahrai, derek@cim.mcgill.ca, McGill University, Montreal, Quebec, Canada; Jaroslav Křivánek, Charles University, Chaos Czech a. s. Prague, Czech Republic, jaroslav.krivaneck@mff.cuni.cz; Wojciech Jarosz, wojciech.k.jarosz@dartmouth.edu, Dartmouth College, Hanover, NH, USA.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM Transactions on Graphics*, <https://doi.org/10.1145/3355089.3356559>.

require numerical estimation. Using an approximation of the optical thickness, even when unbiased, produces a biased transmittance estimate after exponentiation [Raab et al. 2008].

Instead, unbiased transmittance estimation can be constructed by augmenting the medium with fictitious, “null-scattering” particles, altering its density so as to admit analytic distance sampling between successive photon-particle collisions. An estimate of the transmittance between any two points can then be obtained by initiating a random walk along the connecting segment from one end, and measuring its weight (a.k.a. score) at the other end [Spanier 1966; Woodcock et al. 1965; Krivánek et al. 2014; Novák et al. 2018]. While unbiasedness can be proven here by showing their expected score equals the desired transmittance [Coleman 1968], the efficiency of the resulting estimators is understood from a mostly empirical perspective. Moreover, the deviation from traditional Monte Carlo integration makes the comparative analysis between estimators difficult. The adherence to the null-scattering particle interpretation has limited the options for improving existing estimators to the realm of physical plausibility. Of note, Galtier et al. [2013] reformulated null-scattering algorithms as estimators of a modified radiative transfer equation (RTE) [Chandrasekhar 1960], and Kutz et al. [2017] built atop this formulation to develop random walks for spectrally-varying media. While these works establish a more concrete integral formulation of the problem, they still only treat estimators with design decisions restricted by the aforementioned null-scattered particle interpretation.

We present several novel integral formulations of medium transmittance that, unlike the traditional exponential formulation, are amenable to direct, unbiased Monte Carlo estimation. While our formulations are derived from the RTE, they do not necessarily have any physical interpretation, yet provide new insight into the structure of the problem. Purely mathematical interpretation of light transport, divorced from physical analogies, has previously lead to the development of completely new algorithms, including bidirectional path tracing [Veach and Guibas 1994; Lafortune and Willem 1993] and Metropolis light transport [Veach and Guibas 1997]. We show how taking a similar stance, in the context of the transmittance problem, leads to new and powerful unbiased estimators.

We first show (Section 3) that existing unbiased estimators can all be viewed as direct Monte Carlo estimators of a Volterra integral equation of the second kind [Gripenberg et al. 1990]. It is known that the rendering equation [Kajiya 1986; Immel et al. 1986] is a Fredholm integral equation of the second kind, and Volterra integral equations are similar to those but with variable integration domains. This connection allows us to draw analogies between estimators for the rendering equation and estimators for our Volterra integral equation for transmittance. Moreover, we show that under this formulation, the null-scattering concept is not strictly necessary for unbiased transmittance estimation, yet it can be interpreted as a purely mathematical manipulation for reducing variance of the integrand (Section 3.3).

We then derive three entirely new integral formulations (Section 4) from a power-series expansion of the classical transmittance exponential: one reminiscent of the recursive rendering equation [Kajiya 1986; Immel et al. 1986], another reminiscent of its iterative Neumann series expansion [Arvo 1995], and one analogous

to its (hypercube) path integral formulation [Veach 1997; Kelemen et al. 2002]. These base formulations can be directly estimated using Monte Carlo and optionally manipulated mathematically to interpret null scattering as a control-variate transformation (Section 4.2).

Importantly, by deriving the integral formulations (Section 4) separately from any particular estimation strategy (Section 5), we can formally reason about what properties are intrinsic to the integral problem, and how particular choices during Monte Carlo estimation impact variance and performance. This leads in turn to several novel transmittance estimators with efficiency tradeoffs distinct and complementary to current estimators; Fig. 1 highlights two of our estimators. Finally, our formulations provide a clear recipe for applying sample stratification to further reduce estimation variance.

2 BACKGROUND AND PREVIOUS WORK

We first review the physical law of light attenuation along straight lines and summarize prior research efforts in computing the transmittance between two points.

As light travels a distance t along a ray with origin \mathbf{o} and direction ω through a medium, it collides at location $\mathbf{x} = \mathbf{o} + t\omega$ with particles that either absorb or out-scatter some of its energy. In media with particles distributed statistically independently from each other, the change in radiance L along a differential distance dt is quantified by a reduced version of the differential *radiative transfer equation* (RTE) [Chandrasekhar 1960] that does not account for radiance gains due to self-emission and in-scattering at \mathbf{x} :

$$\frac{dL(\mathbf{o} + t\omega)}{dt} = -\mu(\mathbf{o} + t\omega)L(\mathbf{o} + t\omega), \quad \text{or} \quad \frac{dL(t)}{dt} = -\mu(t)L(t), \quad (1)$$

where the spatially varying *extinction coefficient* $\mu(\mathbf{x})$ gives the probability density for such collisions to occur per unit distance at \mathbf{x} .

In rendering, light transport is often simulated in direction opposing the natural flow of light (see Fig. 2). Making the change of variable $t \equiv -x$ and writing L and μ as functions of the distance x along this opposing direction, i.e. $L(x) \equiv L(\mathbf{o} - x\omega)$ and $\mu(x) \equiv \mu(\mathbf{o} - x\omega)$, yields a variant of the reduced RTE (1) atop which we base all remaining formulations in this paper:

$$-\frac{dL(x)}{dx} = -\mu(x)L(x). \quad (2)$$

2.1 Exponential formulation

To obtain an expression for the transmittance $T(a, b) = L(a)/L(b)$, i.e. the fraction of radiance at point b reaching point $a < b$ (see Fig. 2, top), one can solve Eq. (2) for L , yielding the well-known expression (see Appendix A, also notation in Table 1)

$$T(a, b) = e^{-\int_a^b \mu(x) dx} = e^{-\tau(a, b)}. \quad (3)$$

Obtaining an unbiased estimate of T is difficult in general heterogeneous media, where the *optical thickness* integral τ is not available in analytic form. The classical approach to estimate τ alone, e.g. via ray marching [Perlin and Hoffert 1989], even if done in an unbiased way [Pauly et al. 2000], yields a biased estimate of T after exponentiation [Raab et al. 2008; Novák et al. 2018].

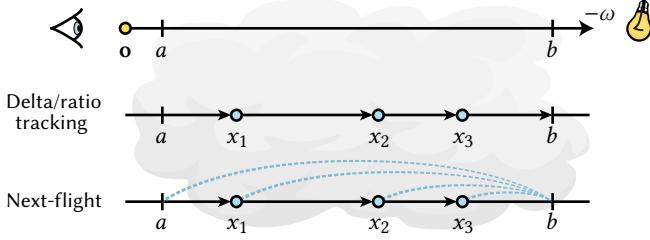


Fig. 2. Illustration of the problem setup (top), the delta- and ratio-tracking estimators (middle), and the next-flight estimator (bottom). Note that estimation is typically performed in direction opposite of light flow.

2.2 Null-scattering estimators

Existing unbiased transmittance estimators are all based on the concept of *null scattering*. Here, one augments the medium with fictitious, perfectly forward-scattering matter with spatially varying extinction μ_n chosen to produce a combined *majorant* medium with simpler, e.g. constant, extinction $\bar{\mu}(x) = \mu(x) + \mu_n(x)$. The free-flight distance $d_i = x_i - x_{i-1}$ between successive collisions along a ray in this augmented medium can then be sampled analytically, with the pdf $p(d_i) = \bar{\mu}(x_i) e^{-\bar{\tau}(x_{i-1}, x_i)}$. Choosing the scattering type at x_i , real or fictitious, according to the respective local albedos, $\mu(x_i)/\bar{\mu}(x_i)$ and $\mu_n(x_i)/\bar{\mu}(x_i)$, respectively, produces an analog random walk, a.k.a. *delta tracking* or Woodcock tracking [Woodcock et al. 1965; Raab et al. 2008]. Below we recapitulate existing transmittance estimators based on this concept.

Track-length estimator. Delta tracking can be used to estimate transmittance between a and b by starting a walk from a and terminating it if real scattering occurs before reaching b . The probability of reaching b without scattering in the real medium is equal to the transmittance $T(a, b)$, motivating the *track-length* estimator name [Spanier 1966; Spanier and Gelbard 1969; Cramer 1978; Raab et al. 2008]. The estimator scores one if b is reached and zero otherwise:

$$\widehat{T}_{tl}(a, b) = \begin{cases} 1 & \text{if } b \leq x_{k+1}, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where x_{k+1} is the first event after k null-scattering events.

Ratio-tracking estimator. Instead of relying on stochastic chance to reach b , the ratio-tracking estimator deterministically chooses to continue forward at each collision with the majorant medium. It then returns the accumulated null-scattering probabilities as a weight [Cramer 1978; Novák et al. 2014]:

$$\widehat{T}_{rt}(a, b) = \prod_{i=1}^k \frac{\bar{\mu}(x_i) - \mu(x_i)}{\bar{\mu}(x_i)} = \prod_{i=1}^k \frac{\mu_n(x_i)}{\bar{\mu}(x_i)}, \quad (5)$$

where x_k is the last collision before reaching b : $x_k < b \leq x_{k+1}$. This estimator has lower variance than the binary track-length estimator but also higher computation cost.

Residual-tracking estimator. Novák et al. [2014] applied control variates to reduce the variance of transmittance estimation, decomposing the extinction μ into a sum of an analytically integrable control component $\underline{\mu}$ and a residual component: $\mu(x) = \underline{\mu}(x) + [\mu(x) - \underline{\mu}(x)]$.

Table 1. List of commonly used notations throughout the document. While occasionally we call $\bar{\mu}$ and $\underline{\mu}$ extinction majorant and minorant, we allow these to be set arbitrarily. The only scene-dependent quantity is μ .

Notation	Description
$\mu(x)$	medium extinction coefficient at distance x
$\bar{\mu}(x), \underline{\mu}(x)$	extinction upper control (e.g. majorant) and lower control (e.g. minorant); free parameters, $\underline{\mu}(x) \leq \bar{\mu}(x)$
$\mu_n(x)$	null extinction coefficient: $\mu_n(x) = \bar{\mu}(x) - \mu(x)$
$\tau, \bar{\tau}, \underline{\tau}, \tau_n$	optical thickness $\int \square(x) dx$ for $\square \in \{\mu, \bar{\mu}, \underline{\mu}, \mu_n\}$; unless given as function arguments, integral bounds are a, b
$T, \widehat{T}, \widehat{\tau}_n$	transmittance, transmittance estimator, τ_n estimator

The control transmittance $e^{-\underline{\tau}}$ is then computed analytically and the residual transmittance $e^{\tau - \underline{\tau}}$ is estimated, e.g. via ratio tracking:

$$\widehat{T}_{rtt}(a, b) = e^{-\underline{\tau}(a, b)} \prod_{i=1}^k \frac{\bar{\mu}(x_i) - \underline{\mu}(x_i)}{\bar{\mu}(x_i) - \mu(x_i)} \quad (6a)$$

$$= e^{-\underline{\tau}(a, b)} \prod_{i=1}^k \frac{\mu_n(x_i)}{\bar{\mu}(x_i) - \underline{\mu}(x_i)}, \quad (6b)$$

where, as with Eq. (5), x_k is the last collision before reaching b . Szirmay-Kalos et al. [2011] proposed a similar decomposition, but instead estimated the residual transmittance using delta tracking (4).

Next-flight estimator. At each step x_i in the ratio-tracking random walk, we can analytically compute the uncollided transport $e^{-\bar{\tau}(x_i, b)}$ to b through the majorant medium, rather than estimating it randomly based on the next free-flight distance. Summing these *next-flight* estimates yields [Cramer 1978]

$$\widehat{T}_{nf}(a, b) = e^{-\bar{\tau}(a, b)} + \sum_{i=1}^k e^{-\bar{\tau}(x_i, b)} \prod_{j=1}^i \frac{\mu_n(x_j)}{\bar{\mu}(x_j)}. \quad (7)$$

This next-flight estimator has only recently been introduced to graphics [Kutz et al. 2017; Novák et al. 2018].

Discussion. The track-length estimator has been originally constructed following physical principles, and the ratio-tracking and next-flight estimators have been derived from it by manipulation that preserves the expected value. Since these estimators are not formalized in a common Monte Carlo integration framework, the understanding of their efficiency has been limited to intuition and numerical experiments. For example, the ratio-tracking and next-flight estimators remain unbiased even when the null extinction μ_n is (locally) negative, even though such configurations are physically implausible. How to choose μ_n , how to improve the efficiency of these estimators, and how to devise new unbiased transmittance estimators have remained open questions. We believe the main reason for this is that the connection between the estimators, e.g. Eq. (7), and the quantity they estimate, Eq. (3), is not obvious.

One of our goals in this paper is to reformulate transmittance estimation in terms of a simple integration problem. To this end, in the following two sections we derive several integral formulations of transmittance from Eqs. (2) and (3) that are amenable to direct and unbiased Monte Carlo estimation.

3 VOLTERRA INTEGRAL FORMULATIONS

Our first transmittance formulation is a recursive integral equation. A base formulation follows almost trivially from Eq. (2). We also derive a null-scattering extension from which the estimators discussed in Section 2.2 arise naturally.

3.1 Base formulation

Integrating Eq. (2) from a to b yields

$$-\int_a^b \frac{dL(x)}{dx} dx = L(a) - L(b) \stackrel{(2)}{=} -\int_a^b \mu(x)L(x)dx \quad (8a)$$

$$L(a) = L(b) - \int_a^b \mu(x)L(x)dx, \quad (8b)$$

where, for clarity, we explicitly denote that the second equality in Eq. (8a) holds due to Eq. (2). Dividing Eq. (8b) by $L(b)$ yields an expression for the transmittance $T(a, b) = L(a)/L(b)$:

$$T(a, b) = 1 - \int_a^b \mu(x)T(x, b) dx. \quad (9)$$

The integral on the RHS can be interpreted as the medium opacity between a and b , i.e. the fraction of collided light, which is absorbed and out-scattered; transmittance is the complement of opacity.

Equation (9) is an inhomogeneous Volterra integral equation of the second kind [Gripenberg et al. 1990], where T is the unknown function (of the first argument; the second argument b is considered a fixed parameter), the source term in front of the integral is one, and the kernel is $-\mu$. Its structure is similar to the well-known surface rendering Fredholm integral equation [Kajiya 1986], with the difference that the integration bounds in Eq. (9) change at every recursion level, as we illustrate in Fig. 3, left.

3.2 Monte Carlo estimation of the base formulation

Due to their similarities, many numerical methods used to solve Fredholm integral equations can be applied to Volterra integral equations. For instance, we can first trivially estimate Eq. (9) by replacing the integral with a one-sample Monte Carlo (MC) estimator at each recursive level on the RHS: $T(a, b) \approx \widehat{T}(a, b) = 1 - \mu(x)\widehat{T}(x, b)/p(x)$, where x is drawn proportionally to some pdf $p(x)$. Applying Russian roulette with continuation probability $P_{\text{rec}} < 1$ to terminate the recursion results in the following unbiased estimator:

$$\widehat{T}(a, b) = 1 - \frac{\mu(x)}{p(x)P_{\text{rec}}} \widehat{T}(x, b). \quad (10)$$

This estimator is akin to a unidirectional path tracer with Russian roulette termination applied to the surface rendering equation. However, unlike that equation where the kernel, i.e. the bidirectional scattering distribution function (BSDF), is normalized to integrate to at most one, here the kernel, $-\mu$, can have arbitrarily high magnitude over the entire integration domain. This makes $\mu(x)/(p(x)P_{\text{rec}})$ susceptible to extreme variation, since $p(x)$ has to be a valid normalized pdf. Equation (10) is thus prone to producing high-variance estimates whose sign will also be a random variable due to the changing sign at each recursion level.

3.3 Null-scattering reinterpretation

We will now manipulate Eq. (2) (prior to integration and MC estimation) in a manner that will afford flexibility for reducing variance, and which is mathematically equivalent to the concept of augmenting the media with fictitious null-scattering matter. We start by introducing a free parameter $\bar{\mu}(x)$, which we will later on call *extinction upper control*, to rewrite the extinction coefficient $\mu(x)$ as

$$\mu(x) = \bar{\mu}(x) - [\bar{\mu}(x) - \mu(x)] = \bar{\mu}(x) - \mu_n(x), \quad (11)$$

which can then be substituted into Eq. (2):

$$-\frac{dL(x)}{dx} = -[\bar{\mu}(x) - \mu_n(x)]L(x). \quad (12)$$

Note that this transformation is valid for any choice of $\bar{\mu}$, even though the physical interpretation of introducing fictitious matter makes sense only when its corresponding “null” density μ_n is non-negative everywhere, i.e. when $\bar{\mu}(x) \geq \mu(x)$. We therefore adhere to a purely mathematical view of Eq. (12) as a manipulation of Eq. (2) that is not conditioned on physical plausibility. This also avoids the need to assign a non-physical delta forward-scattering phase function to the fictitious matter, as done in prior formulations [Galtier et al. 2013; Kutz et al. 2017; Novák et al. 2018].

Equation (12) is a first-order linear differential equation that can be written in canonical form and solved for L using standard techniques. We do that in Appendix B to obtain a null-scattering extension to our base integral formulation (9) (see notation in Table 1):

$$T(a, b) = e^{-\bar{\tau}(a, b)} + \int_a^b e^{-\bar{\tau}(a, x)} \frac{[\bar{\mu}(x) - \mu(x)]}{\mu_n(x)} T(x, b) dx. \quad (13)$$

Note the appearance of two exponential terms, as well as the replacement of the extinction μ by the null extinction μ_n which also absorbs the negative sign in front of the integral.

Recall that $\bar{\mu}$ is a free functional parameter; setting $\bar{\mu}(x) = \mu(x)$ makes the integral vanish, yielding the traditional exponential transmittance formulation (3). Another special case is $\bar{\mu}(x) = 0$, where both exponentials simplify to one, yielding our base Volterra formulation (9). The parameter $\bar{\mu}$ thus defines a continuum of formulations. Importantly, if $\bar{\mu}(x) \geq \mu(x)$ for every $x \in [a, b]$, the kernel $e^{-\bar{\tau}(a, x)}\mu_n(x)$ is strictly non-negative and integrates to at most one (just like the BSDF in the rendering equation), which has implications on the variance of estimators, as we will discuss below.

Note that our formulation in Eq. (13) is equivalent to that of Kutz et al. [2017, Appendix B]. However, theirs is written around a specific class of estimators that use a certain sampling distribution for x . In practice, any valid distribution can be used, and we adhere to a clear distinction between the integral *formulations* and their numerical *estimation*, which we discuss next.

3.4 Monte Carlo estimation of null-scattering formulation

As before, Eq. (13) is a Volterra integral equation. A direct application of Monte Carlo estimates the integral by evaluating the integrand at a location $x \in [a, b]$ drawn from a pdf $p(x)$. The term $T(x, b)$ can be estimated recursively, terminating via Russian roulette with continuation probability P_{rec} . We can additionally apply Russian roulette

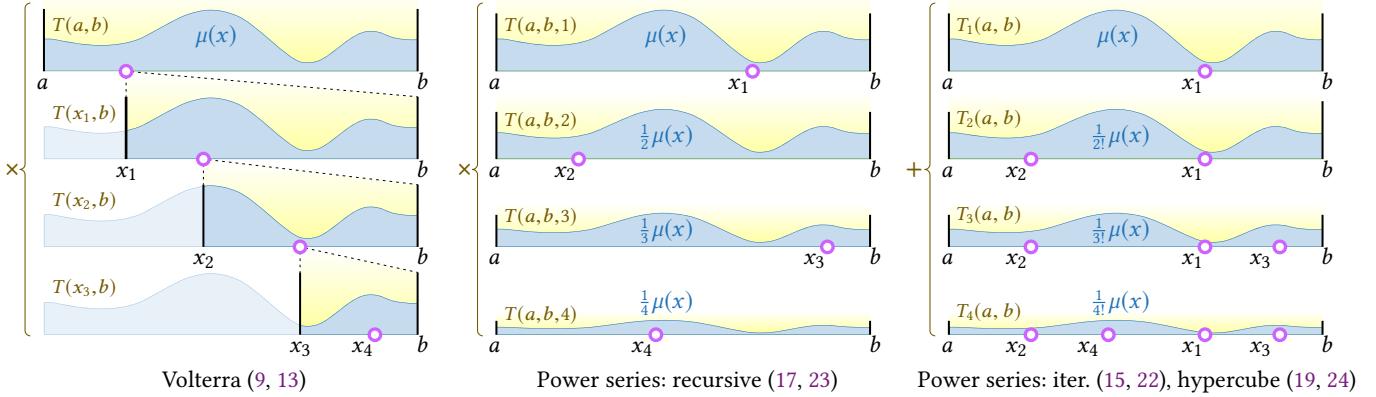


Fig. 3. Graphical illustrations of our formulations for the transmittance $T(a, b)$ and their integration bounds (highlighted in yellow), here showing four terms from each formulation which are either multiplied (left and middle) or summed (right) together. Note how in the Volterra formulations the integration bounds shrink at every recursion level. The power-series based formulations instead maintain the original bounds $[a, b]$ but progressively scale the integral.

to stochastically evaluate the source term $e^{-\bar{\tau}}$ with probability P_{src} . This leads to the following general *unidirectional estimator*:

$$\widehat{T}(a, b) = \frac{e^{-\bar{\tau}(a, b)}}{P_{\text{src}}} + \frac{e^{-\bar{\tau}(a, x)} \mu_n(x) \widehat{T}(x, b)}{p(x) P_{\text{rec}}}. \quad (14)$$

This estimator again bears similarity to a unidirectional path tracer with Russian roulette. Below we discuss how, for specific choices of sampling pdf $p(x)$ and Russian roulette probabilities P_{src} and P_{rec} , this estimator reduces to the four existing estimators we reviewed in Section 2.2 (illustrated schematically in Fig. 2).

3.5 Reduction to existing estimators

We start by noting that a good choice for $p(x)$ is one that importance samples one or more of the integrand's non-recursive terms, e.g. the exponential term. Using $p(x) = \bar{\mu}(x) e^{-\bar{\tau}(a, x)}$ we can also terminate the recursion when x is sampled past b , corresponding to the continuation probability $P_{\text{rec}} = \Pr\{x > b\} = 1 - e^{-\bar{\tau}(a, b)}$. Note that $p(x)$ then needs to be re-normalized over the interval $[a, b]$ with a factor equal to $1/P_{\text{rec}}$. Cancelling out the relevant terms in the rightmost term of Eq. (14) simplifies it to $\mu_n(x)/\bar{\mu}(x)\widehat{T}(x, b)$.

Next-flight estimator. Using the above density $p(x)$ and recursion probability P_{rec} , and always evaluating the source term (via $P_{\text{src}} = 1$), yields the next-flight estimator (7) after expanding the recursion.

Ratio-tracking estimator. Alternatively, we can choose to evaluate exactly one of the two additive terms in Eq. (14) at each recursion level. Setting $P_{\text{src}} = 1 - P_{\text{rec}} = e^{-\bar{\tau}(a, b)}$ correlates the sampling of P_{rec} and P_{src} , making the estimator evaluate the source term if $x > b$ and estimate the integral otherwise. Expanding the recursion yields the ratio-tracking estimator (5).

Track-length estimator. The score of the ratio-tracking estimator is multiplied by the “null”-albedo factor $\mu_n(x)/\bar{\mu}(x)$ at every recursion level. Multiplying P_{rec} by another Russian roulette continuation probability $P_{\text{null}} = \mu_n(x)/\bar{\mu}(x)$ cancels this factor to maintain a running score of one, akin to the albedo-based termination commonly used in path tracing. Expanding the recursion yields the

binary track-length estimator (4). Note that the P_{null} sampling is independent from P_{src} and P_{rec} which are dependent on the integral estimation variable x . Additionally, P_{null} is a valid probability only when $\mu_n(x) \geq 0 \Leftrightarrow \bar{\mu}(x) \geq \mu(x)$, i.e. when the physical interpretation of null scattering is valid.

Residual-tracking estimator. Equation (6) has been originally formulated as analytical control-transmittance evaluation coupled with residual-transmittance estimation. In our framework, we can interpret that entire equation as a direct MC estimator for Eq. (13), just like ratio tracking but with a different pdf $p(x)$. This follows from the observation that distance sampling in the residual medium is driven by the residual majorant $\bar{\mu} - \underline{\mu}$, i.e. $p(x) = [\bar{\mu}(x) - \underline{\mu}(x)] e^{-[\bar{\tau}(a, x) - \underline{\tau}(a, x)]}$. Plugging this pdf into Eq. (14), adjusting the probabilities $P_{\text{src}} = 1 - P_{\text{rec}} = e^{-[\bar{\tau}(a, b) - \underline{\tau}(a, b)]}$, simplifying, and expanding the recursion yields Eq. (6b). Thus, the only difference to regular ratio tracking (5) is the use of a mean sampling distance of $1/(\bar{\mu} - \underline{\mu})$ instead of $1/\bar{\mu}$.

The concept of residual estimation has also been applied to the track-length estimator [Novák et al. 2014] and suggested for the next-flight estimator [Novák et al. 2018]. In our framework these correspond to plugging the above residual-majorant exponential pdf into the track-length and next-flight variants of estimator (14).

3.6 Discussion

Eq. (14) provides a framework to help interpret transmittance estimators, mindfully separating the impact of the underlying integrand, sampling distribution, and continuation probabilities. We show that certain sampling choices yield known estimators. Importantly, setting $\bar{\mu}(x) \geq \mu(x)$ bounds their estimates between 0 and 1, avoiding potentially high variance in the base Volterra estimator Eq. (10). We call the free parameter $\bar{\mu}$ the extinction upper control due to its effectiveness when bounding the extinction μ .

Beyond that, however, it remains difficult to reason about the efficiency of existing estimators. To date, the modern understanding of efficiency is built empirically and can sometimes lead to counter-intuitive behavior. For example, next-flight estimation uses Rao-Blackwellization [Blackwell 1947] to substitute a random decision—the conditional evaluation of the source term $e^{-\bar{\tau}}$ —by its expected

value, and so one would expect it to consistently outperform ratio-tracking estimators. This does not hold, as experiments in prior work have shown [Novák et al. 2018]. Another example is the inconsistent impact that $\bar{\mu}$ has on different estimators: increasing $\bar{\mu}$ reduces the average sampled distance (i.e. estimators take smaller steps), but this only improves ratio-tracking estimators and not next-flight estimators, as also observed in earlier work [Novák et al. 2018]. Finally, residual tracking samples distances according to a distribution that produces larger steps on average, yet it never under-performs ratio tracking [Novák et al. 2014].

These apparent inconsistencies hint to a gap in the current understanding. The Monte Carlo estimates we have explored treat the recursive nature of the integration problem *implicitly*, focusing on minimizing the variance of sub-estimates at a fixed recursion level using *local* stochastic decisions (e.g. drawing samples proportionally to a local density, continuation probabilities). Here, we again draw an analogue between our Volterra formulation and the surface rendering equation, both of which provide local uni-dimensional views of an infinite-dimensional integration problem. For the rendering equation, a more global view of the problem falls from an integral reparameterization onto the space of full light transport paths. This opens up opportunities to extend beyond incremental estimation with local sampling techniques (e.g. BSDF importance sampling) to more global techniques (e.g. bidirectional sampling [Veach and Guibas 1994; Lafourcade and Willem 1993]). We will present a transmittance formulation that similarly informs more global interpretations of our problem, and so too new estimators.

4 POWER-SERIES FORMULATIONS

In this section we present three novel integral formulations of transmittance derived from a power-series expansion of the classical exponential formulation (3). This will allow us to write transmittance analogously to the rendering path integral, and to interpret the null-scattering transformation as an application of control variates. This, in turn, will enable us to more easily apply and analyze new Monte Carlo estimation techniques to this problem (Section 5).

4.1 Base formulations

While the classical formulation (3) includes an integral, its exponentiation prevents the direct application of MC integration for unbiased transmittance estimation. To this end, our idea is to transform that exponential to make it amenable to direct MC estimation.

Iterative formulation. Applying the standard power-series expansion of the exponential function, $e^{-\tau} = \sum_{k=0}^{\infty} (-\tau)^k / k!$, and writing out the optical thickness integral τ , yields

$$T(a, b) = \sum_{k=0}^{\infty} T_k(a, b) = \sum_{k=0}^{\infty} \frac{(-\tau)^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \prod_{i=1}^k \left[-\int_a^b \mu(x) dx \right]. \quad (15)$$

This equation expresses transmittance as an infinite sum of integral products. Similar power-series expansions have been applied to other integral transformations, e.g. reciprocals [Booth 2007] to eliminate the bias in the photon mapping algorithm [Qin et al. 2015].

Note that the terms T_k resemble the probability mass function (pmf) of a Poisson distribution $\Pr\{X = k\} = e^{-\lambda} \lambda^k / k!$ with mean

$\lambda = \tau$, only unnormalized and with alternating signs for the even and odd k . This insight will later help us to gain a better understanding of the efficiency of existing transmittance estimators.

Recursive formulation. Expanding the sum in Eq. (15) and regrouping the terms gives

$$T(a, b) = \sum_{k=0}^{\infty} \prod_{i=1}^k \left(-\frac{\tau}{i} \right) = 1 - \frac{\tau}{1} \left[1 - \frac{\tau}{2} \left[1 - \frac{\tau}{3} [\dots] \right] \right]. \quad (16)$$

Extracting the following recurrence relation from the RHS of Eq. (16):

$$T(a, b, i) = 1 - \frac{\tau}{i} T(a, b, i+1) = 1 - \frac{1}{i} T(a, b, i+1) \int_a^b \mu(x) dx, \quad (17)$$

the transmittance between two points is given by $T(a, b) = T(a, b, 1)$. Note that unlike Eq. (9), the integration bounds here do not change at every recursion level. Each level therefore computes the same integral, and this expansion of the integration bounds is compensated for by the scaling factor $1/i$. We illustrate this in Fig. 3, middle.

Hypercube formulation. Every summand T_k in Eq. (15) can be written as a k -dimensional integral:

$$T(a, b) = \sum_{k=0}^{\infty} \frac{1}{k!} \underbrace{\int_a^b \cdots \int_a^b}_{k \text{ times}} \left[\prod_{i=1}^k -\mu(x_i) \right] dx_1 \cdots dx_k. \quad (18)$$

Each such integral runs over a k -hypercube $\mathcal{H}_k = [a, b]^k$. Their sum can be merged into one integral over the union $\mathcal{H} = \bigcup_{k=0}^{\infty} \mathcal{H}_k$:

$$T(a, b) = \sum_{k=0}^{\infty} \int_{\mathcal{H}_k} f_k(\mathbf{x}) dx = \int_{\mathcal{H}} f(\mathbf{x}) dx, \quad (19)$$

where $\mathbf{x} = (x_1, \dots, x_k) \in \mathcal{H}_k$ for $k \geq 0$,¹ and the integrand $f(\mathbf{x})$ and volume measure dx are defined for each k separately as

$$f(\mathbf{x}) = \frac{1}{k!} \prod_{i=1}^k -\mu(x_i), \quad dx = \prod_{i=1}^k dx_i. \quad (20)$$

Equation (19) expresses transmittance as a pure integration problem, with the unknown T appearing only on the LHS, where the integral runs over the union of all hypercubes $[a, b]^k$. We illustrate this formulation in Fig. 3, right.

Analogy to rendering formulations. Along with our Volterra formulation (9), our recursive power-series formulation (17) is reminiscent of the rendering equation [Kajiya 1986], and its iterative expansion (15) is the counterpart to the Neumann series expansion of that equation [Arvo 1995]. Our hypercube formulation (19) bears similarity to the hypercube path integral formulations of light transport of Kelemen et al. [2002]. Similarly to that formulation, it enables the application of general variance reduction techniques, such as the multiple importance sampling (MIS) combination of different sampling techniques [Veach and Guibas 1995], which we will demonstrate in Section 5. In Appendix C we show how Eqs. (15) and (19) can instead be derived from our base Volterra formulation (9).

¹Note that for $k = 0$, the zero-hypercube \mathcal{H}_0 is a single point, thus the integral $\int_{\mathcal{H}_0} f(\mathbf{x}) dx$ simplifies to $f(\cdot) = 1/0! = 1$, which is indeed the 0th summand in Eq. (15).

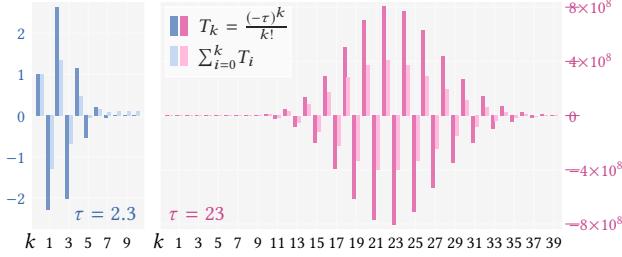


Fig. 4. Plots of the first few terms T_k in Eq. (15) for two values of τ , along with their prefix sums (in lighter colors) which converge to the respective transmittance values, $T = 0.1$ and $T = 0.1 \times 10^{-9}$, at infinity. Note the large magnitudes of individual terms relative to the small transmittance values.

Discussion. The above formulations express transmittance as an infinite summation problem. In practice this sum can only be estimated by evaluating a finite number of terms, and we will discuss specific techniques in Section 5. The efficiency of such estimation will be directly influenced by the values of these terms, and it is therefore useful to study their shape. In Fig. 4 we plot the first few terms T_k of Eq. (15) for two values of τ . We also plot the prefix sums which converge to the transmittance T as $k \rightarrow \infty$. In accordance with our earlier observation, each plot resembles the pmf of a Poisson distribution with mean τ , only that it is unnormalized and the terms have alternating signs. More importantly, note the high dynamic range in the plots: individual terms can have extremely large magnitudes, positive and negative, though these cancel out when summed up to a transmittance value in the range $[0, 1]$. The mode (i.e. peak) of the Poisson-like distribution is $(-\tau)^{[x]}/[\tau]!$ and is located at the mean index $k = \lfloor \tau \rfloor$. Since τ can be any non-negative number, the terms of the transmittance sum can have arbitrary magnitude. Estimators for this sum are thus susceptible to extreme variance, especially if they are unaware of the location of high-magnitude terms.

The cause of the extreme magnitude variation in the above base formulations is the presence of negative terms T_k in the underlying infinite sum. If we could transform this sum to ensure each term is non-negative, then each one would have to be bounded between zero and one for the sum to converge to a valid transmittance value. This is what null-scattering approaches effectively do, and in the next subsection we show how this transformation can be viewed as a control variate on the base integral formulations.

4.2 Null-scattering formulations as control variates

To derive the null-scattering extensions of our power series integral formulations, we begin by substituting the transformation from Eq. (11) into the exponential transmittance expression (3):

$$T(a, b) = e^{-\int_a^b \mu(x) dx} = e^{-\int_a^b \bar{\mu}(x) dx} e^{\int_a^b [\bar{\mu}(x) - \mu(x)] dx} = e^{-\bar{\tau}} e^{\tau_n}. \quad (21)$$

Control variates use an analytically integrable control function, which is correlated with the original integrand, to reduce variance during Monte Carlo estimation. Here the task is to compute the integral $-\tau = -\int_a^b \mu(x) dx$ but we have another function $\bar{\mu}(x)$ whose integral $\bar{\tau}$ we can obtain analytically. Adding $\bar{\mu}(x)$ to our original integrand and also subtracting $\bar{\tau}$ keeps the problem unchanged since

$-\tau = \int_a^b [\bar{\mu}(x) - \mu(x)] dx - \bar{\tau}$; however, subsequent MC estimation of the remaining integral will have reduced variance if the functions μ and $\bar{\mu}$ are positively correlated. Note that this control-variate transformation is identical to that of Novák et al. [2014, Sec. 3.2], though instead of $\bar{\mu}$ theirs uses an extinction lower control $\underline{\mu}$ parameter (e.g. set to the extinction minorant). We will discuss how their method maps to our framework in more detail in Section 5.1.

Iterative formulation. We can expand the term e^{τ_n} in Eq. (21) into a power series to obtain a control-variate extension of Eq. (15):

$$T(a, b) = \sum_{k=0}^{\infty} \frac{e^{-\bar{\tau}} \tau_n^k}{k!} = e^{-\bar{\tau}} \sum_{k=0}^{\infty} \frac{1}{k!} \prod_{i=0}^k \int_a^b \frac{[\bar{\mu}(x) - \mu(x)]}{\mu_n(x)} dx. \quad (22)$$

Just like in our null-scattering Volterra formulation (13), the negative sign in front of the integral is now absorbed by the integrand $\mu_n(x)$, which is non-negative whenever $\bar{\mu}(x) \geq \mu(x)$. Moreover, note that the terms T_k are proportional to the terms $e^{-\tau_n} \tau_n^k/k!$ of a Poisson pmf with mean τ_n . (The terms are exactly equal when $\tau_n = \bar{\tau} \Leftrightarrow \tau = 0$, i.e. when $T(a, b) = 1$.) This insight is useful when constructing estimators for this formulation, as we will see in Section 5.

Recursive formulation. The recursive form of Eq. (22) is derived analogously to Eq. (17):

$$T(a, b, i) = e^{-\bar{\tau}} + \frac{1}{i} T(a, b, i+1) \overbrace{\int_a^b \frac{[\bar{\mu}(x) - \mu(x)]}{\mu_n(x)} dx}^{\tau_n}. \quad (23)$$

As before, the transmittance is given by $T(a, b) = T(a, b, 1)$.

Hypercube formulation. The control-variate extension to our hypercube formulation (19) can be derived from Eq. (22) by following the steps we took in Eqs. (18) and (19), obtaining

$$T(a, b) = \int_{\mathcal{H}} f_n(\mathbf{x}) d\mathbf{x}. \quad (24)$$

The volume measure $d\mathbf{x}$ is the same as in Eq. (20) and the contribution function $f_n(\mathbf{x})$ is now defined (for each k separately) as

$$f_n(\mathbf{x}) = e^{-\bar{\tau}} \frac{1}{k!} \prod_{i=1}^k \mu_n(x_i).. \quad (25)$$

Discussion. The formulations in Eqs. (22) to (24) extend their base counterparts from Section 4.1 by introducing the free extinction upper control parameter $\bar{\mu}$. Recall that the purpose of transforming the sum in Eq. (15) was to ameliorate the extreme variation of its terms. To that end, it is interesting to study the effect of the control optical thickness $\bar{\tau} = \int_a^b \bar{\mu}(x) dx$ on Eq. (22). In Fig. 5 we plot the terms T_k and their prefix sums for a fixed real optical thickness $\tau = 2.3$ and four different control thicknesses $\bar{\tau}$. Expectedly, all four prefix sums converge to the same transmittance value $T = e^{-2.3} = 0.1$, however their shapes are heavily affected by $\bar{\tau}$. When $\bar{\tau} = 0$, Eq. (22) simplifies to Eq. (15); the result is plotted Fig. 4, left. When $0 \leq \bar{\tau} < \tau$ (Fig. 5, far left), the magnitude of the terms is reduced but negative values still remain. Another special case is $\bar{\tau} = \tau$ (Fig. 5,

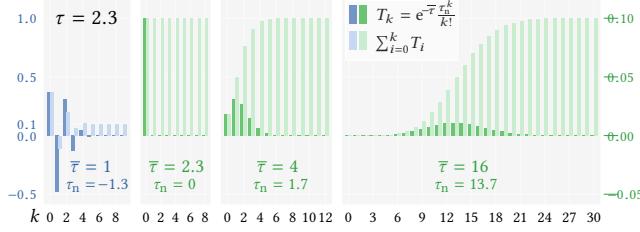


Fig. 5. Plots of the first few terms T_k of Eq. (22) and their prefix sums (in lighter colors) for $\tau = 2.3$ and four different values of $\bar{\tau}$. Note that all prefix sums converge to the same transmittance value $T = e^{-2.3} = 0.1$. For $\bar{\tau} \geq \tau$ (green plots), individual terms are always in the range $[0, 1]$; prefix-sum convergence is then monotone, with rate depending on $\bar{\tau}$. Notice that the terms T_k form an unnormalized Poisson pmf with mean $\tau_n = \bar{\tau} - \tau$, with their prefix sum being the corresponding CMF. The normalization factor for this to be a valid pmf is the inverse of the transmittance T .

middle left), where $\tau_n = 0$ and all “mass” is concentrated in the first term T_0 whose value equals the transmittance T .

When $\bar{\tau} \geq \tau$ (Fig. 5, middle right and far right), the terms T_k in Eq. (22) form a proper Poisson pmf, only unnormalized. Individual values are bounded between 0 and 1, and their prefix sum is the corresponding (unnormalized) Poisson cumulative mass function (CMF). The normalization factor is the inverse of the total sum, i.e. the inverse of the transmittance $e^{-\tau}$. The mode of this pmf (i.e. its maximum value) is $e^{-\bar{\tau}} \tau_n^{\lfloor \tau_n \rfloor} / \lfloor \tau_n \rfloor!$, located at the mean index $k = \lfloor \tau_n \rfloor$, with $\tau_n = \bar{\tau} - \tau$. Note that as $\bar{\tau}$ increases, the mode decreases but also moves further to the right. This reduces the variation in the pmf, but at the same time pushes the center of mass to the right and makes the CMF converge more slowly.

In summary, for $\bar{\tau} > 0$, the values of the terms T_k in Eq. (22) always vary less than those in Eq. (15). The control-variate (i.e. null-scattering) transformation thus effectively tames the variation in the base formulations, bounding it to be between 0 and 1 for $\bar{\tau} \geq \tau$.

5 POWER SERIES ESTIMATION

In this section, we explore the Monte Carlo estimation of our three control-variate formulations from Section 4.2, noting that they simplify to their base counterparts from Section 4.1 when $\bar{p}(x) = 0$.

We first reformulate the existing random-walk based estimators (Section 2.2) in our power-series framework, which reveals new insight into their efficiency. We then propose several novel estimators tailored to our formulations, and show how different estimators can be combined via MIS. Finally, we discuss how to correctly apply sample stratification to further improve the estimators’ efficiency, which we will analyze further in Section 6.

5.1 Single-term estimation

We begin by addressing our iterative power-series formulation (22). The simplest way to estimate this infinite sum is to evaluate a single term $T_k = e^{-\bar{\tau}} \tau_n^k / k!$ with index k chosen with probability $P(k)$. The value of T_k needs to be estimated too as it contains an integral:

$$\widehat{T}(a, b) = \frac{\widehat{T}_k(a, b)}{P(k)} = \frac{e^{-\bar{\tau}}}{k! P(k)} \prod_{i=1}^k \widehat{\tau}_{n,i} = \frac{e^{-\bar{\tau}}}{k! P(k)} \prod_{i=1}^k \frac{\mu_n(x_i)}{p(x_i)}. \quad (26)$$

Here, $\widehat{\tau}_{n,i} = \mu_n(x_i)/p(x_i)$, for $i = 1, \dots, k$, are k independent estimates of the null thickness integral τ_n using samples $x_i \in [a, b]$ with some density $p(x_i)$. Note that while each of the samples x_i is used to estimate the same integral, they all have to be mutually uncorrelated in order for the expectation of their product to be equal to τ_n^k .

Ratio-tracking, residual-tracking, and track-length estimators. To evaluate Eq. (26), one way to distribute uncorrelated points $x_i \in [a, b]$ is through a Poisson point process. One way to realize this is by sampling the distances between neighboring points from an exponential distribution [Ross 1996]. We observe that this, in fact, is precisely how the random-walk estimators reviewed in Section 2.2 operate, using the exponential distance pdf $p(x_i - x_{i-1}) = \bar{p}(x_i) e^{-\bar{\tau}(x_{i-1}, x_i)}$. The number k of generated samples $x_i \in [a, b]$ cannot be explicitly controlled with this method, but it is a discrete random variable that follows a Poisson distribution with mean $\bar{\tau}$. This is a result of the Poisson point process, and so is the density of the samples [Ross 1996]. Using the above distance pdf, these are respectively

$$P(k) = e^{-\bar{\tau}} \frac{\bar{\tau}^k}{k!} \quad \text{and} \quad p(x) = \frac{\bar{p}(x)}{\bar{\tau}}. \quad (27)$$

Substituting $P(k)$ and $p(x_i)$ into Eq. (26), all terms cancel out except for $\mu_n(x_i)$ and $\bar{p}(x_i)$, arriving at the ratio-tracking estimator (5). Additionally applying Russian roulette termination with probability $\mu_n(x_i)/\bar{p}(x_i)$ at every x_i yields the track-length estimator (4). Analogously to ratio tracking, residual tracking (6) can be viewed as a Poisson point process with Poisson pmf $P(k)$ with mean $\bar{\tau} - \underline{\tau}$ (instead of $\bar{\tau}$) and sample density $p(x) = (\bar{p}(x) - \underline{p}(x)) / (\bar{\tau} - \underline{\tau})$.

P-series ratio/residual-tracking estimators. Having mapped ratio tracking to our power-series framework, an alternative novel implementation of that estimator arises: Instead of having the number of steps k being the result of a Poisson point process, as above, we can first choose the index k of the summand to estimate, with Poisson probability $P(k)$, and then generate k independent samples x_i in the interval $[a, b]$ with pdf $\bar{p}(x_i)/\bar{\tau}$. These samples are used to construct k independent one-sample estimates $\widehat{\tau}_{n,i}$ of the null optical thickness. This approach produces samples with the same density as the random-walk based ratio-tracking estimator, and thus the same estimator (5). However it uses a different sampling procedure that generates the points x_i in an arbitrary order unlike the traditional Poisson point process implementation where $x_i \leq x_j$ for $i < j$. We will refer to it as our *p*-series ratio-tracking estimator. We can analogously create a power-series variant of residual tracking by explicitly sampling from its corresponding pmf $P(k)$ and pdf $p(x)$.

Discussion. Recall that the summands of Eq. (22) are proportional to a Poisson pmf with mean τ_n . That pmf is thus ideal for single-term estimation (26), however its parameter $\tau_n = \bar{\tau} - \tau$ is unknown (since τ is unknown). An approximation can be used instead: the closer it is to τ_n , the lower the variance of the estimator.

Ratio tracking uses a Poisson sampling pmf with mean $\bar{\tau}$, which has the property $\lim_{\bar{\tau} \rightarrow \infty} \tau_n = \bar{\tau}$, since $\tau_n = \bar{\tau} - \tau$. The estimator thus becomes better at importance sampling the transmittance sum as

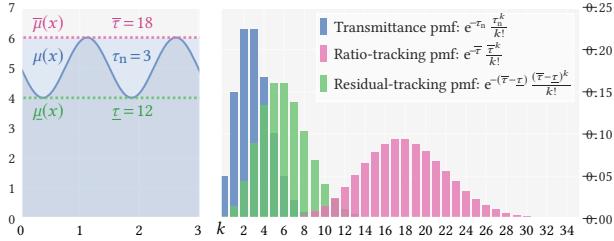


Fig. 6. Single-term estimation of the sum in Eq. (22) (blue bars, normalized for display to a Poisson pmf with mean $\tau_n = 3$) for the extinction function μ on the left. Ratio tracking selects one term using a Poisson pmf with mean $\bar{\tau} = 18$. Residual tracking uses a pmf with mean $\bar{\tau} - \underline{\tau} = 6$. Once chosen, the cost of estimating the term (in terms of medium lookups) is proportional to its index. Note that the residual-tracking pmf is more closely proportional to the ideal (blue) pmf than the ratio-tracking pmf is. It thus achieves better importance sampling, also at a lower cost due to its smaller pmf mean.

the extinction upper control $\bar{\mu}$ increases, which is in line with prior empirical observations [Novák et al. 2014, 2018]. Residual tracking instead uses a pmf with mean $\bar{\tau} - \underline{\tau}$, which can be a much closer approximation to τ_n when $\bar{\mu}$ and $\underline{\mu}$ tightly bound the variation of the extinction μ , respectively from above and below. We illustrate this in Fig. 6, where we plot the transmittance terms (normalized to a pmf for display) for a sinusoidal extinction function, along with the sampling pmfs of ratio and residual tracking. Note that the green residual pmf is more closely proportional to the blue transmittance terms than the pink ratio pmf is. Moreover, the average cost of using the residual pmf is lower, since on average it chooses terms with lower index k , and estimation cost is proportional to k in terms of the number of medium extinction lookups $\mu_n(x_i) = \bar{\mu}(x_i) - \mu(x_i)$.

In Appendix D we derive analytical expressions for the variance and the efficiency of the ratio-tracking, track-length, and residual-tracking estimators as functions of τ and $\bar{\tau}$, for the case where τ_n is estimated with zero variance, i.e. when $\bar{\mu}(x) \propto \mu(x)$.

5.2 Prefix-sum estimation: iterative

Another way to estimate the infinite sum in Eq. (22) is to evaluate not just a single term T_k , but all terms up to and including T_k :

$$\widehat{T}(a, b) = \sum_{j=0}^k \frac{T_j(a, b)}{\Pr\{k \geq j\}} = \sum_{j=0}^k \frac{e^{-\bar{\tau}} \prod_{i=1}^j \widehat{\tau}_{n_i}}{j! [1 - C(j)]} = \sum_{j=0}^k \frac{e^{-\bar{\tau}} \prod_{i=1}^j \frac{\mu_n(x_i)}{ip(x_i)}}{j! [1 - C(j)]}, \quad (28)$$

where the index k follows some pmf $P(k)$, and $C(k) = \sum_{j=0}^{k-1} P(j)$ is the corresponding cumulative mass function (CMF). Thus $1 - C(j)$ above is the probability of $k \geq j$, i.e. of evaluating term j . As in Eq. (26), $x_i \sim p(x_i)$ are k independent samples for estimating τ_n , and a prefix x_1, \dots, x_j of these are used to evaluate each term j . Thus the cost of this estimator is identical to that of Eq. (26), even though it evaluates k terms instead of just one.

P-series next-flight estimator. To evaluate the prefix-sum estimator (28), we can use the distributions from Eq. (27). We will refer to this variant as our *p-series next-flight* estimator since it is similar to the next-flight estimator (7), which evaluates a prefix-sum for an expansion of our Volterra formulation (13) instead.

5.3 Prefix-sum estimation: recursive

We can also construct a prefix-sum estimator for our power-series formulation (22) based on its recursive form (23):

$$\widehat{T}(a, b, i) = e^{-\bar{\tau}} + \frac{\mu_n(x_i)}{ip(x_i)P_i} \widehat{T}(a, b, i+1), \quad (29)$$

where $x_i \in [a, b]$ is a random sample with pdf $p(x_i)$ used to estimate the null-thickness integral τ_n at every level i of the recursion which is continued with probability P_i . We propose two estimators based on Eq. (29), both using the sample pdf $p(x) = \bar{\mu}(x)/\bar{\tau}$ from Eq. (27) and differing only in the choice of continuation probabilities P_i .

P-series cumulative estimator. This estimator accumulates the weights w_i during the recursion and uses continuation probabilities

$$P_i = \min(|W_i|, 1) = \min\left(\left|\frac{\mu_n(x_i)}{ip(x_i)} \prod_{j=0}^{i-1} w_j\right|, 1\right), \quad (30)$$

which try to maintain a cumulative weight $W_i = 1$ during the recursion. Since the weight can be negative when $\mu_n(x_i) < 0$, we take its absolute value. The same style of Russian roulette termination is often used in path tracing, where the continuation probability is set proportionally to the throughput along the path [Pharr et al. 2016].

P-series CMF estimator. To achieve low-variance prefix-sum estimation, we can take advantage of the knowledge that the terms of the power-series sum (22) form an (unnormalized) Poisson pmf (see Fig. 5). We can start applying Russian roulette termination only after we have accumulated a certain fraction t of the corresponding (unnormalized) CMF. The Poisson parameter τ_n is unknown (i.e. to be estimated), but we can use the free parameter $\bar{\tau}$ as a proxy. We define the continuation probability at recursion level i as

$$P_i = \begin{cases} \bar{\tau}/i & \text{if } C_{\bar{\tau}}(i) \geq t, \\ 1 & \text{otherwise,} \end{cases} \quad (31)$$

where $C_{\bar{\tau}}$ is the CMF of the Poisson pmf with mean $\bar{\tau}$. We set $t = 0.99$, and use $\bar{\tau}/i$ as a proxy for the expected weight $\tau_n/i = E\left[\frac{\mu_n(x_i)}{ip(x_i)}\right]$ at level i . When $\bar{\tau} \geq \tau$, the delayed termination guarantees that the accuracy of the estimate will be at least t (99% in our case). Unlike the above cumulative estimator, the cost of this estimator is only dependent on $\bar{\mu}$ (via $\bar{\tau}$), and not on μ , just like ratio tracking.

5.4 Hypercube estimation

A primary (i.e. one-sample) estimator for the pure integral transmittance formulation in Eq. (24) has the well-known general form

$$\widehat{T}(a, b) = \frac{f_n(\mathbf{x})}{p(\mathbf{x})}, \quad (32)$$

where $\mathbf{x} = (x_1, \dots, x_k)$ is a point in the k -hypercube \mathcal{H}_k sampled with density $p(\mathbf{x})$. In this formulation, the only difference between estimators can be due to the pdfs $p(\mathbf{x})$ of their corresponding sampling techniques. Inspired by bidirectional path tracing in global light transport, we explore a family of techniques, each denoted as (k, t) that constructs $\mathbf{x} = (x_1, \dots, x_t, x_{t+1}, \dots, x_k)$, where the first $t \in [0..k]$ coordinates are sampled via a random walk starting from

a and the remaining $k - t$ coordinates starting from b . Denoting $x_0 \equiv a$ and $x_{k+1} \equiv b$, the pdf for such a technique is

$$p_{k,t}(\mathbf{x}) = \frac{1}{k!} \overbrace{\prod_{i=1}^t \bar{\mu}(x_i) e^{-\bar{\tau}(x_{i-1}, x_i)}}^{\text{from } a} \overbrace{\prod_{i=t+1}^k \bar{\mu}(x_i) e^{-\bar{\tau}(x_i, x_{i+1})}}^{\text{from } b} \quad (33a)$$

$$= \frac{1}{k!} e^{-\bar{\tau}(a, x_t) - \bar{\tau}(x_{t+1}, b)} \prod_{i=1}^k \bar{\mu}(x_i). \quad (33b)$$

This technique samples the coordinates of \mathbf{x} in a sorted order, i.e. in a k -simplex of the hypercube \mathcal{H}_k , with volume $|\mathcal{H}_k|/k!$. The factor $1/k!$ is a Jacobian that converts the pdf measure to that of the hypercube.

Hypercube ratio-tracking and next-flight estimators. Ratio tracking corresponds to using the technique $(k, k+1)$; writing out the estimator (32) with $p_{k,k+1}$ in the denominator yields Eq. (5) after canceling out the relevant terms. The next-flight estimator uses techniques (i, i) , for $i = 0, \dots, k$, to construct a set of points $\mathbf{x}_i = (x_1, \dots, x_i)$ with pdfs $p_{i,i}(\mathbf{x}_i)$ that share common coordinate prefixes. Summing up corresponding estimates (32) yields Eq. (7).

Unidirectional MIS estimator. Our hypercube transmittance formulation allows us to combine the ratio-tracking and next-flight estimators via MIS, just like path tracing combines unidirectional sampling and next-event estimation. This starts a random walk from a , and at every step x_i constructs a next-flight estimate with pdf $p_{i,i}(\mathbf{x}_i)$, which is weighted by a balance-heuristic weight [Veach and Guibas 1995] that also accounts for the corresponding ratio-tracking technique with pdf $p_{i,i+1}(\mathbf{x}_i)$. Upon reaching b , a ratio-tracking estimate with pdf $p_{k,k+1}(\mathbf{x}_k)$ is accumulated, where the MIS weight also accounts for the next-flight technique with pdf $p_{k,k}(\mathbf{x}_k)$.

Bidirectional MIS estimator. We can also combine all techniques (k, t) for every point \mathbf{x} via MIS, similarly to how bidirectional path tracing (BPT) combines all possible vertex connection techniques for a given light transport path [Veach and Guibas 1995]. Our implementation samples coordinates by starting two random walks—one from a and one from b , and then constructs points \mathbf{x} by concatenating coordinate prefixes, analogously to a standard BPT implementation that connects pairs of vertices on sensor and emitted paths.

5.5 Sample stratification

Due to the sequential sample placement in prior formulations, it has not been clear how stratification can be correctly applied to improve the efficiency of the transmittance estimation without introducing bias. Sample stratification is best understood in an integration framework, which is what our formulations provide. Every term T_k in the power-series sum (22) contains a product of k one-dimensional optical thickness integrals. Estimates of these integrals have to be uncorrelated so that the expectation of their product is the product of their expectations, τ_n^k . Multiple estimation samples within each of these integrals, however, *can* be correlated, i.e. stratified.

In practice, our estimators take one sample per integral, but we stratify samples across estimation calls, i.e. across the N paths traced for an individual image pixel. Since the power-series sum is infinite, we need to be able to construct the N -sample pattern on-the-fly as its dimensionality can be arbitrarily high depending on the index

k of the term being estimated. One option is to use quasi-Monte Carlo patterns, such as Hammersley [Wong et al. 1997]. Noting that individual integrals are one-dimensional, we can also utilize on-the-fly Latin-hypercube/ N -rooks sampling [McKay et al. 1979; Shirley 1991], which ensures stratification along each 1D projection. We also stratify the discrete sampling decisions in our estimators, i.e. the choice of index k and Russian roulette termination. Effectively, each sampling decision, continuous or discrete, consumes one dimension of the sampling pattern. This logic can be also applied to stratify the distance and termination decisions in the Volterra estimators.

6 SUMMARY AND RESULTS

In this section we recapitulate our transmittance formulations and the various estimators we discussed above. We then benchmark these estimators on a range of scenes, summarizing the results as simple guidelines for which estimators to use in different conditions.

6.1 Estimator summary

Our Volterra formulations (Section 3) express transmittance as an integral equation. Existing techniques (Section 2.2) arise as direct, recursive Monte Carlo estimators for that formulation (Section 3.5).

Our power-series formulations (Section 4) express the same quantity as an infinite sum. We explore three types of estimators for that sum. The existing track-length, ratio-tracking, and residual-tracking estimators are *single-term* estimators (Section 5.1). We propose novel power-series variants of these, also of the existing next-flight estimator which can be viewed as a *prefix-sum* estimator (Section 5.2). We additionally devise two novel recursive prefix-sum estimators that only differ in their termination probabilities: p-series cumulative and p-series CMF (Section 5.3). Existing estimators can also be expressed as sampling techniques in our *hypercube* formulation; we propose two novel estimators that combine such techniques via MIS, unidirectionally and bidirectionally, respectively (Section 5.4).

The ideal discrete distribution for single-term power-series estimation is the Poisson pmf with mean τ_n . In practice, close approximations of the unknown τ_n will yield low variance. Unfortunately, no ideal target termination pmf exists for prefix-sum estimators, whose variance drops as more terms are accumulated, approaching zero only in the limit where the estimator never terminates. Low-variance estimation can still be achieved by delaying the termination until after the majority of the contribution has been accumulated, i.e. after the prefix sum has almost converged to the sought transmittance value (see the plots in Fig. 5 for intuition). This of course comes at increased cost for evaluating a larger number of terms.

6.2 Numerical experiments

We benchmark the efficiency of the various estimators on both canonical configurations and real scenes. To compare only transmittance estimation, we use a path tracer that follows all surface reflection and refraction branches and avoids all forms of Russian roulette external to evaluating transmittance. In our rendering comparisons, we hold the number of medium extinction lookups constant and report the root mean squared error (RMSE) and render time of the images. Equal-lookup comparisons are renderer-agnostic and arguably more representative for large, complex scenes, where medium lookups dominate the render time [Novák et al. 2014, 2018].

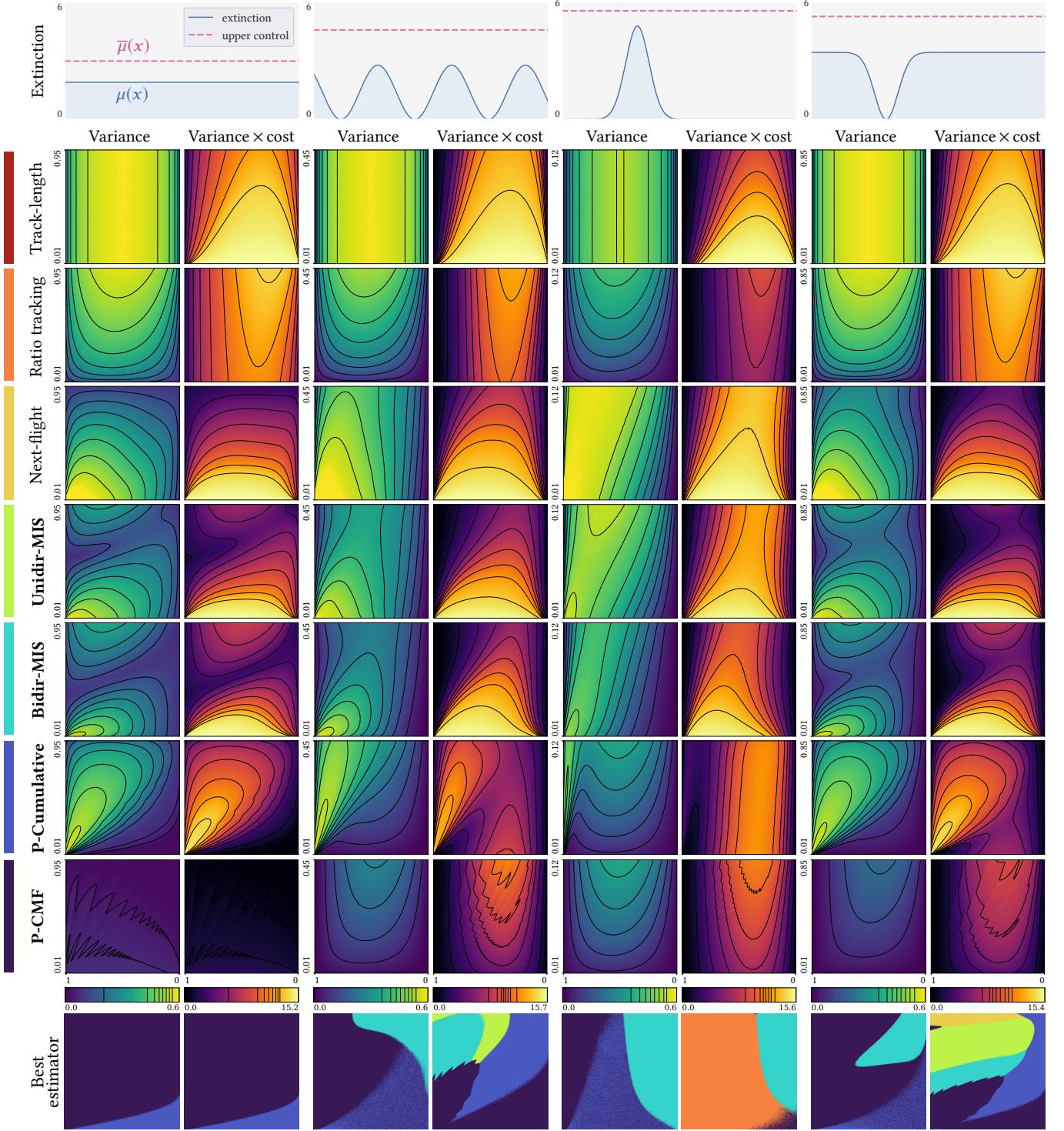


Fig. 7. Canonical comparison of estimators (ours are in bold) on several extinction functions. Each pixel in every 2D plot represents a transmittance estimation experiment with a unique combination of extinction function μ and (constant) extinction upper control $\bar{\mu}$. In each plot, we scale μ exponentially along the x-axis so that the transmittance drops linearly from 1 to 0. Along the y-axis we scale $\bar{\mu}$ so that the control thickness efficiency $\tau/\bar{\tau}$ increases linearly from 0.01 to an extinction-dependent upper bound. For each pixel, we run the estimator several times to compute its variance and cost (i.e. number of μ lookups). We then plot the variance and cost \times variance (i.e. inverse efficiency), where lower is better. In the bottom-row plots we show the best estimator for every pixel in color code. Our p-series CMF estimator performs best overall, in terms of both metrics, thanks to trading slightly higher cost for lower variance. A notable exception is the Gaussian extinction function, where the upper control is naturally loose and the less costly ratio-tracking estimator performs marginally better.

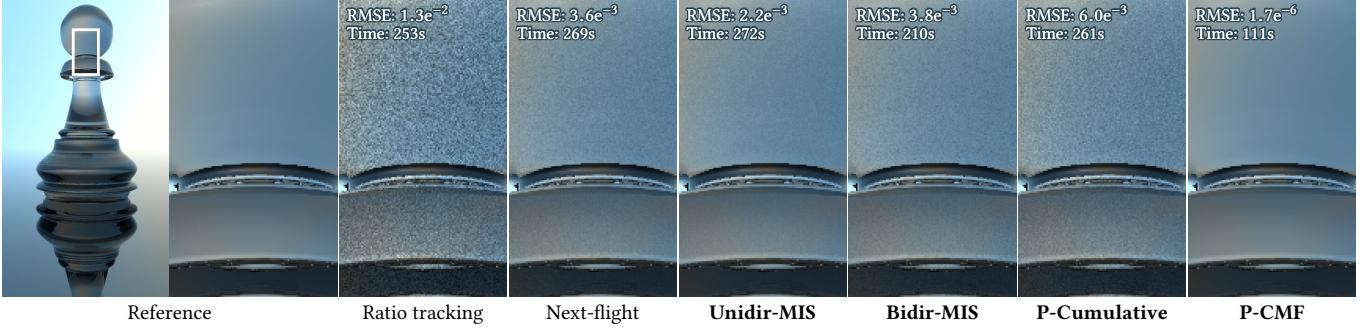


Fig. 8. Equal-cost (92.5M extinction lookups) comparison of several estimators (ours are in bold) on a homogeneous medium with extinction upper control set to $\bar{\mu} = 1.5 \times \mu$. The relatively tight $\bar{\mu}$ favors next-flight and MIS estimators over ratio tracking. Our p-series CMF estimator produces a virtually noise-free result; its render time is also low since lookups are cheap in this scene and this estimator makes more lookups per ray than the others.

Canonical configuration. In Fig. 7 we study the performance of seven estimators across four extinction functions in a canonical configuration previously used by Novák et al. [2014, 2018]. We plot the estimators’ variance and variance \times cost (i.e. inverse efficiency), where lower is better. In each 2D plot, we scale the extinction function μ such that the transmittance to be estimated varies linearly from 1 to 0, left to right. Vertically, we scale the (constant) extinction upper control $\bar{\mu}$ such that the control thickness efficiency $\tau/\bar{\tau}$ [Lepänen 2010] increases linearly from 0.01 to an extinction-dependent upper bound, bottom up. The plots convey the relative differences in performance between estimators on the full range of possible transmittance values and (bounding) constant upper controls: from tight (top) to almost infinitely loose (bottom).

The performance of existing estimators (top three rows in Fig. 7) has been studied experimentally in prior work [Novák et al. 2014, 2018].² Our power-series reformulations (Section 5.1) provide a more rigorous understanding of their behavior. As the upper control increases, ratio tracking’s Poisson pmf becomes better at importance sampling the transmittance sum, making up for the increased cost to improve the estimator’s overall efficiency. The track-length estimator’s variance is never smaller than that of ratio tracking, but its cost can be lower to result in higher efficiency, as is the case when the optical thickness is large and the upper control is tight (top right regions of the plots). These observations match what our analytical efficiency expressions for these estimators in Appendix D suggest. The next-flight estimator has been known to perform better when the upper control is tight [Novák et al. 2018]. In such cases, the mean of its termination Poisson pmf is larger than that of the (unnormalized) transmittance pmf, so this prefix-sum estimator terminates after accumulating the majority of the transmittance contribution. For loose controls, termination comes too early; when infinitely loose, the two pmf means coincide and on average only half the contribution is accumulated. The estimator is thus bound to produce high variance when constant controls are naturally loose, e.g. with Gaussian-shaped extinctions.

Our p-series cumulative prefix-sum estimator is designed to delay termination until its cumulative weight W_i (30) (i.e. the analog of

path throughput in surface rendering) drops below one. This typically occurs well past the transmittance pmf mean, resulting in better performance than the next-flight estimator for the majority of cases. The estimator is still susceptible to some higher variance when the extinction control is loose (i.e. the thickness efficiency $\tau/\bar{\tau}$ is low) and also the null thickness $\tau_n = \bar{\tau} - \tau$ is small, e.g. with values $\bar{\tau} = 0.7$ and $\tau_n = 0.5$ (bottom left plot regions). In this case the true transmittance is $e^{-\tau} = e^{-0.2} \approx 0.82$, and the estimator terminates at the first iteration with probability $P_1 = W_1 \approx \tau_n$, i.e. half the time, scoring $e^{-\tau} \approx 0.5$; otherwise it terminates almost surely at the second iteration, scoring $2e^{-\tau} \approx 1$.

Our p-series CMF estimator mitigates such high variance by making sure to suppress Russian roulette termination until almost all of the contribution has been accumulated. This allows it to have the lowest variance among all estimators. It also has the highest efficiency, except for the Gaussian extinction where ratio tracking performs slightly better due to p-series CMF being more conservative, i.e. trading higher cost for lower variance.

Our Unidir-MIS is the MIS combination of ratio-tracking and next-flight estimators enabled by our theory, and Bidir-MIS additionally includes other sampling strategies (see Section 5.4). MIS improves robustness by eliminating the worst-case behaviors of any individual strategy, but it also does not attain the best-case behavior of any strategy in isolation.

Homogeneous medium. In media with constant density, such as the one in Fig. 8, the null-thickness integral τ_n is estimated with zero variance when using a constant extinction upper control. Thus the only source of variance in the transmittance estimators is the discrete random choice of terms/recursion termination in their corresponding power-series formulations.

The upper control is relatively tight in this scene, so ratio tracking performs poorly as its Poisson pmf does not importance sample the transmittance sum well. Next-flight does better than ratio tracking and also p-series cumulative, though terminates relatively early and some noise remains. P-series CMF is more conservative and provides near-perfect estimation. It also does so in a fraction of the time as it makes more lookups per ray, and lookups are cheap in this simple medium. The benefits of combining techniques via MIS suggested by the plots in Fig. 7 can also be seen visually here.

²Note that the ‘variance’ and ‘variance \times cost’ plots of Novák et al. [2014, Fig. 9] and Novák et al. [2018, Fig. 5] incorrectly use the estimators’ standard deviation instead of their variance, hence the discrepancies between their plots and ours.

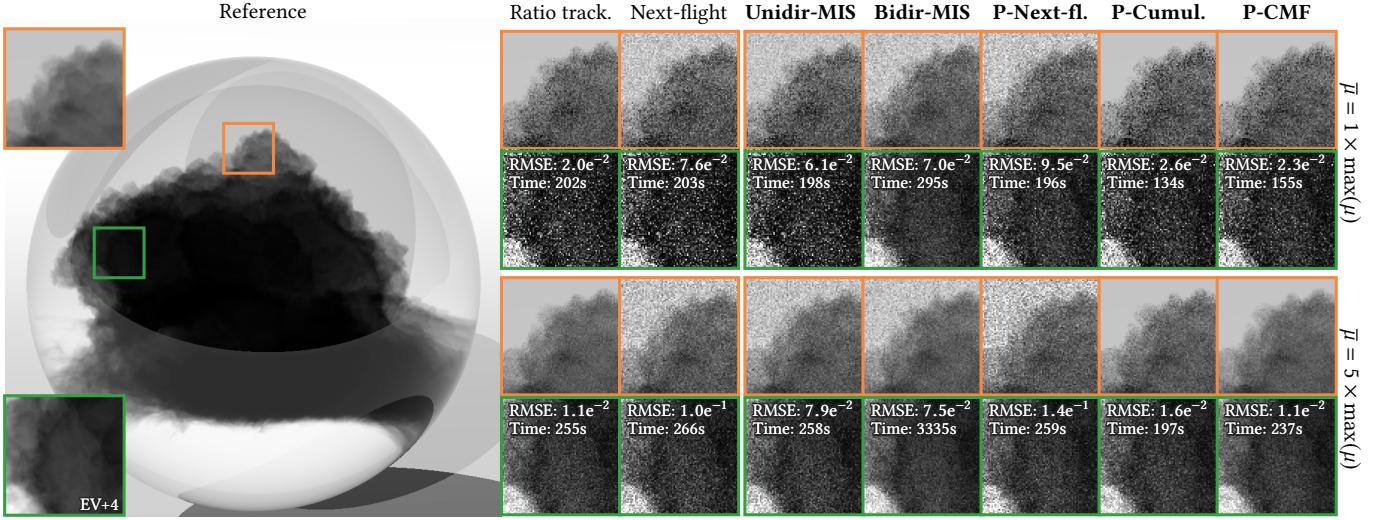


Fig. 9. Equal-cost (0.5B and 1.3B lookups, top and bottom respectively) comparisons of estimators (ours are in bold) on a heterogeneous medium for two upper controls. Even the tighter control (top rows) is relatively loose, favoring ratio tracking. Our p-series cumulative and CMF estimators do almost as well.

Heterogeneous medium. We now consider the more complex heterogeneous medium in Fig. 9 and compare the estimators’ performance for two constant upper controls $\bar{\mu}$: one tight (top rows), equal to the medium’s majorant extinction $\max(\mu)$, and one loose (bottom rows), equal to $5 \times \max(\mu)$. The extinction along rays through the globe resembles that of the Gaussian function in Fig. 7. The tighter $\bar{\mu}$ thus still produces a loose control thickness $\bar{\tau}$, which is a good case for ratio tracking. P-series CMF has a slightly worse RMSE but is faster in terms of render speed as it evaluates fewer samples per pixel for an equal number of extinction lookups. As the upper control loosens, it closes the gap to ratio tracking. P-series cumulative comes a close third. The next-flight and MIS estimators perform relatively poorly due to the next-flight techniques terminating too early. Estimating a transmittance value of 1 with a loose upper control is particularly difficult for such estimators, as can be seen in the orange zoom-ins and also in the bottom left corners of the plots in Fig. 7.

Despite its high overall error, Bidir-MIS performs well in the dense, low-transmittance region of the medium (top green zoom-in in Fig. 9), which corresponds to the right part of the estimator’s plots in Fig. 7. However, not scale well to loose upper controls as the number of techniques it needs to evaluate and weigh is proportional to the upper control thickness. This is reflected by the drastically increased render time in the bottom green zoom-in in Fig. 9. Nevertheless, it provides motivation for future work in attempting to use different weighting schemes or combining other estimators.

Non-bounding upper control. In practice, it can be difficult or costly to compute a bounding (i.e. majorant) upper control, e.g. if the medium density is modified procedurally. In addition, the time required to evaluate transmittance using any known or proposed estimators depends proportionally on the control value. So for sparse media, where only a few small regions reach the maximum density, it would be desirable for an estimator to support non-bounding controls to reduce render time.

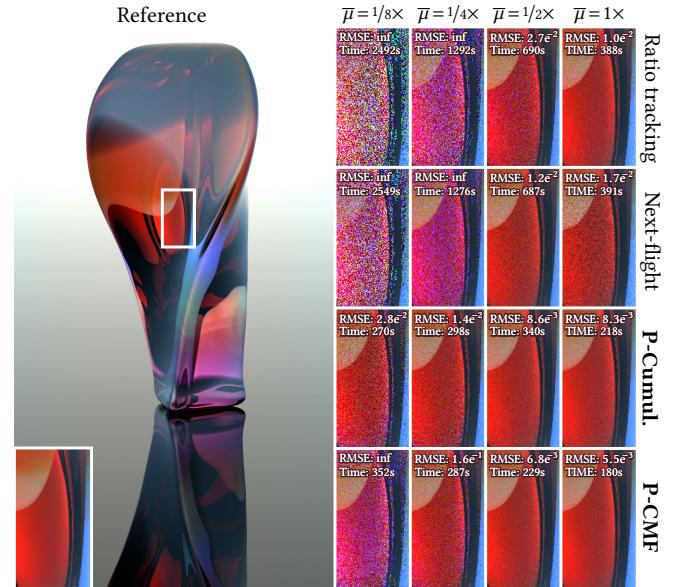


Fig. 10. Equal-cost (160M lookups) render comparison, where in each column we set the extinction upper control $\bar{\mu}$ to a different fraction of the medium’s majorant extinction, $\max(\mu)$. Our p-series estimators (in bold) are more resilient to non-bounding controls than the state of the art.

In Fig. 10 we compare our p-series estimators against the current state of the art when utilizing varying degrees of non-bounding controls. The render times and variance of the existing estimators explode, even reaching numerical infinity for a handful of pixels. For ratio tracking, this aligns with the result in Eq. (44) which suggests the variance grows exponentially with decreasing controls.

Referring back to the leftmost plot in Fig. 5, non-bounding controls produce high-magnitude terms with alternating signs. The

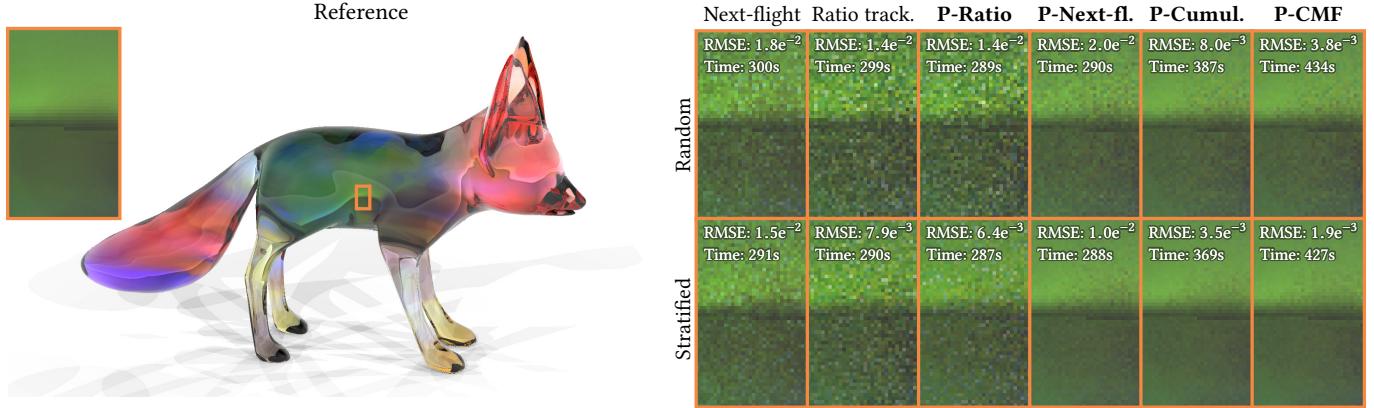


Fig. 11. Comparing the benefits from Latin-hypercube sample stratification that various estimators (ours are in bold) get, using 8 samples/pixel. The power-series ratio-tracking and next-flight variants benefit more from stratification than their traditional (Volterra) counterparts.

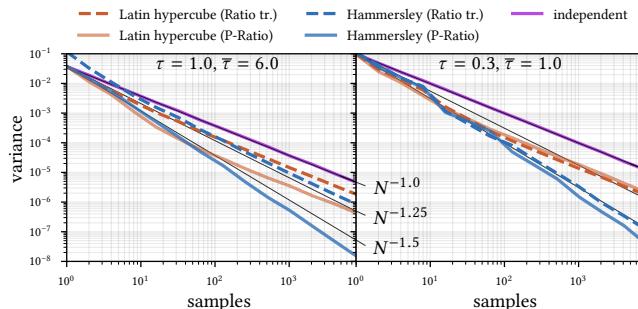


Fig. 12. Convergence plots of ratio tracking vs. our p-series ratio tracking using various sampling methods on a cosine extinction function for two choices of $\bar{\tau}$ and τ . Stratification benefits are dependent on both parameters.

prefix sum of these terms also has significant variation initially, but after accumulating enough terms the oscillation stops and the sum starts approaching the transmittance value. The next-flight estimator performs poorly as it always terminates too early. Our p-series prefix-sum estimators degrade much more gracefully. The cumulative estimator effectively suppresses Russian roulette termination until the absolute value its cumulative weight drops below 1. The CMF estimator does slightly worse as its termination is directly driven by the upper control: when the control is low, the estimator can terminate too early and its variance can explode.

Stratification. In Fig. 11 we show the benefit of sample stratification, notably comparing the traditional Volterra formulations of the ratio-tracking and next-flight estimators and our p-series variants. Expectedly, the two variants of ratio tracking perform identically with independent sampling, but our p-series formulation benefits more from stratification. In our variant, the discrete choice of term T_k is explicit and easily stratified, thereby significantly reducing variance. The plots in Fig. 12 also show an improvement in the variance convergence rate. The amount of improvement depends on the configuration, and the Hammersley point set consistently

Table 2. RMSE numbers for the rendered scenes above, where in each row the extinction upper control $\bar{\mu}$ set to a different scale of the medium majorant extinction, $\max(\mu)$. Cell colors are interpolated per row between the best (green) and worst (red) estimator. Black means numerical infinity or not applicable (e.g. track-length does not work with non-bounding controls).

$\bar{\mu}$	Track-length	Ratio track.	Next-flight	P-Ratio	P-Next-fl.	Unidir-MIS	Bidir-MIS	P-Cumul.	P-CMF
Fig. 8	0.1	2.1×10^{-1}	2.8×10^{-1}	8.8×10^{-1}	1.8×10^{-1}	7.3×10^{-1}	9.6×10^{-2}	2.3×10^{-2}	8.8×10^{-2}
	0.3	5.4×10^{-2}	7.3×10^{-2}	3.4×10^{-2}	8.1×10^{-2}	1.6×10^{-2}	1.1×10^{-2}	8.0×10^{-3}	2.1×10^{-3}
	0.6	8.6×10^{-3}	8.2×10^{-3}	5.0×10^{-3}	8.9×10^{-3}	4.6×10^{-3}	6.9×10^{-3}	6.0×10^{-3}	
	1.0	8.2×10^{-3}	1.3×10^{-2}	2.8×10^{-2}	1.9×10^{-2}	2.9×10^{-2}	2.2×10^{-3}	5.8×10^{-3}	6.9×10^{-3}
	1.5	8.2×10^{-3}	1.3×10^{-2}	2.8×10^{-2}	1.9×10^{-2}	2.9×10^{-2}	4.5×10^{-3}	6.1×10^{-3}	4.5×10^{-3}
	2.5	7.8×10^{-3}	8.8×10^{-3}	6.1×10^{-3}	8.9×10^{-3}	4.9×10^{-3}	1.8×10^{-2}	6.2×10^{-3}	7.9×10^{-3}
	5.0	7.8×10^{-3}	5.6×10^{-3}	7.9×10^{-3}	5.6×10^{-3}	6.7×10^{-3}	6.2×10^{-3}	7.1×10^{-3}	2.9×10^{-3}
Fig. 9	0.1	3.4×10^0	2.9×10^0	4.4×10^0	2.7×10^0	8.6×10^0	1.7×10^1	1.7×10^0	2.5×10^0
	0.3	2.2×10^{-2}	1.8×10^{-2}	2.4×10^{-2}	1.7×10^{-2}	1.3×10^{-2}	2.0×10^{-2}	1.1×10^{-2}	8.9×10^{-3}
	0.6	1.0×10^{-2}	1.0×10^{-2}	1.0×10^{-2}	1.9×10^{-2}	1.2×10^{-2}	2.1×10^{-2}	8.3×10^{-3}	5.5×10^{-3}
	1.0	1.2×10^{-2}	7.4×10^{-3}	1.8×10^{-2}	7.4×10^{-3}	1.3×10^{-2}	9.9×10^{-3}	6.1×10^{-3}	9.9×10^{-3}
	1.5	1.2×10^{-2}	5.1×10^{-3}	1.8×10^{-2}	5.1×10^{-3}	1.3×10^{-2}	1.8×10^{-2}	4.0×10^{-3}	2.8×10^{-3}
	2.5	1.1×10^{-2}	3.4×10^{-3}	1.8×10^{-2}	3.4×10^{-3}	1.3×10^{-2}	1.8×10^{-2}	2.4×10^{-3}	1.8×10^{-3}
	5.0	1.1×10^{-2}	1.8×10^{-2}	2.4×10^{-2}					
Fig. 10	0.1	8.9×10^{-2}	3.5×10^{-2}	1.4×10^{-1}	3.6×10^{-2}	4.9×10^{-2}	2.7×10^{-2}	1.6×10^{-2}	1.7×10^{-2}
	0.3	2.4×10^{-2}	1.9×10^{-2}	2.4×10^{-2}	2.0×10^{-2}	1.6×10^{-2}	2.3×10^{-2}	1.4×10^{-2}	1.1×10^{-2}
	0.6	1.4×10^{-2}	1.5×10^{-2}	2.0×10^{-2}	1.5×10^{-2}	2.1×10^{-2}	1.4×10^{-2}	2.3×10^{-2}	1.3×10^{-2}
	1.0	1.7×10^{-2}	1.4×10^{-2}	2.5×10^{-2}	1.4×10^{-2}	2.8×10^{-2}	1.8×10^{-2}	2.0×10^{-2}	1.2×10^{-2}
	1.5	1.6×10^{-2}	9.6×10^{-3}	2.5×10^{-2}	3.0×10^{-2}	1.9×10^{-2}	2.0×10^{-2}	6.7×10^{-3}	4.3×10^{-3}
	2.5	1.6×10^{-2}	9.6×10^{-3}	2.5×10^{-2}	3.0×10^{-2}	2.0×10^{-2}	2.5×10^{-2}	3.8×10^{-3}	2.7×10^{-3}
	5.0	1.6×10^{-2}	6.5×10^{-3}	2.6×10^{-2}	6.6×10^{-3}	3.2×10^{-2}	2.0×10^{-2}	2.5×10^{-2}	
Fig. 11	0.1	3.7×10^{-1}	3.7×10^{-1}	3.7×10^{-1}	2.3×10^{-1}	3.5×10^{-1}	2.9×10^{-1}	1.4×10^{-1}	1.4×10^{-1}
	0.3	3.3×10^{-2}	8.2×10^{-2}	3.3×10^{-2}	7.7×10^{-2}	6.7×10^{-2}	8.4×10^{-2}	3.8×10^{-2}	3.6×10^{-2}
	0.6	1.9×10^{-2}	2.0×10^{-2}	7.6×10^{-2}	2.0×10^{-2}	6.1×10^{-2}	7.0×10^{-2}	2.6×10^{-2}	2.5×10^{-2}
	1.0	1.9×10^{-2}	1.5×10^{-2}	7.4×10^{-2}	1.5×10^{-2}	5.9×10^{-2}	6.3×10^{-2}	2.2×10^{-2}	1.7×10^{-2}
	1.5	1.9×10^{-2}	1.1×10^{-2}	7.3×10^{-2}	1.1×10^{-2}	5.6×10^{-2}	5.7×10^{-2}	1.7×10^{-2}	1.2×10^{-2}
	2.5	1.9×10^{-2}	1.1×10^{-2}	7.3×10^{-2}	1.1×10^{-2}	5.4×10^{-2}	7.9×10^{-2}	1.5×10^{-2}	
	5.0	2.7×10^{-2}	1.1×10^{-2}	1.0×10^{-2}	1.4×10^{-2}				

outperforms Latin-hypercube sampling. Our other novel p-series estimators in Fig. 11 benefit from stratification as well, roughly halving the error on average.

6.3 Discussion

In summary, for scenes containing large, sparse media with loose upper controls, ratio tracking is the preferred estimator if the extinction lookups dominate the render time (Fig. 9). When stratification is available, our p-series variant should be used instead (Figs. 11 and 12). Our p-series cumulative estimator is the only one whose variance does not explode with non-bounding upper controls (Fig. 10). Our p-series CMF is the preferred estimator in all other cases, as suggested by the best-estimator plots at the bottom of Fig. 7. Figure 9 also shows that p-series CMF can be faster than ratio tracking in terms of render time, and Table 2 indicates its performance will approach that of ratio tracking as the extinction upper control becomes looser.

7 CONCLUSION

Our integral formulations lead to both better theoretical understandings of the transmittance problem and more effective numerical estimators. We showed that the existing, seemingly very different, estimators can be reformulated as Monte Carlo estimators of a Volterra integral equation. This reformulation allows us to concretely reason about the discrepancies between their expected and actual (relative) performance. It also reveals that the commonly used null-scattering concept is not strictly necessary for numerical estimation and can be viewed as a simple mathematical manipulation.

We then presented three novel formulations—iterative power-series, recursive power-series, and hypercube—whose key benefit is their amenability to direct and unbiased Monte Carlo estimation, along with an interpretation of null scattering as a control-variate transformation for variance reduction. These formulations also bear similarity to different forms of the surface rendering equation, and this analogy leads to several novel estimators with distinct performance characteristics. The additional benefit of direct Monte Carlo estimation is that it enables the effective application of sample stratification.

While our framework establishes a solid mathematical foundation of the problem of estimating medium transmittance, we believe we have only scratched the surface of a plethora of different estimators that are now possible by our formulations. Future work could explore better importance sampling techniques for optical thickness estimation and power-series prefix-sum termination. Incorporating our formulations into a path integral formulation of global light transport [Miller et al. 2019] is also of practical importance. Our interpretation of the null-scattering concept as control variates implies that it is theoretically possible to remove the concept of majorant/control media altogether which can be challenging to define in some practical cases. We also believe that our formulations can be generalized to handle correlated media [Jarabo et al. 2018; Bitterli et al. 2018] by deriving corresponding formulations.

ACKNOWLEDGMENTS

All images have been rendered using PBRT [Pharr et al. 2016]. The chessboard and pieces in Figs. 1 and 8 are from TurboSquid. The OpenVDB [Museth 2013] cloud model in Fig. 9 is from Disney Research. The Fox model in Fig. 11 is from Thingiverse.

This work was supported by a gift from Autodesk, National Science Foundation Grant IIS-1812796, Czech Science Foundation Grant 19-07626S, and Charles University Grant SVV-2017-260452.

REFERENCES

- James Richard Arvo. 1995. *Analytic Methods for Simulated Light Transport*. Ph.D. Thesis. Yale University, New Haven, Connecticut.
- Benedikt Bitterli, Srinath Ravichandran, Thomas Müller, Magnus Wrenninge, Jan Novák, Steve Marschner, and Wojciech Jarosz. 2018. A Radiative Transfer Framework for Non-Exponential Media. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 37, 6 (Nov. 2018), 225:1–225:17. <https://doi.org/10/gfz2cm>
- David Blackwell. 1947. Conditional Expectation and Unbiased Sequential Estimation. *The Annals of Mathematical Statistics* 18, 1 (03 1947), 105–110. <https://doi.org/10/cjr9wk>
- T. E. Booth. 2007. Unbiased Monte Carlo Estimation of the Reciprocal of an Integral. *Nuclear Science and Engineering* 156, 3 (2007), 403–407. <https://doi.org/10/gfzq76>
- Eva Cerezo, Frederic Pérez, Xavier Pueyo, Francisco J. Serón, and François X. Sillion. 2005. A Survey on Participating Media Rendering Techniques. *The Visual Computer* 21, 5 (June 2005), 303–328. <https://doi.org/10/cjxqdt>
- Subrahmanyan Chandrasekhar. 1960. *Radiative Transfer*. Dover Publications, NY.
- W. A. Coleman. 1968. Mathematical Verification of a Certain Monte Carlo Sampling Technique and Applications of the Technique to Radiation Transport Problems. *Nuclear Science and Engineering* 32, 1 (April 1968), 76–81. <https://doi.org/10/gfzndg>
- S. N. Cramer. 1978. Application of the Fictitious Scattering Radiation Transport Model for Deep-Penetration Monte Carlo Calculations. *Nuclear Science and Engineering* 65, 2 (1978), 237–253. <https://doi.org/10/gfzq74>
- M. Galtier, S. Blanco, C. Caliot, C. Coustet, J. Dauchet, M. El Hafi, V. Eymet, R. Fournier, J. Gautrais, A. Khuong, B. Piaud, and G. Terrée. 2013. Integral Formulation of Null-Collision Monte Carlo Algorithms. *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (Aug. 2013), 57–68. <https://doi.org/10/f446pg>
- G. Gripenberg, Stig-Olof. Londen, and Olof J. Staffans. 1990. *Volterra Integral and Functional Equations*. Cambridge University Press.
- David S. Immel, Michael F. Cohen, and Donald P. Greenberg. 1986. A Radiosity Method for Non-Diffuse Environments. *Computer Graphics (Proceedings of SIGGRAPH)* 20, 4 (Aug. 1986), 133–142. <https://doi.org/10/dmjm9t>
- Adrian Jarabo, Carlos Aliaga, and Diego Gutierrez. 2018. A Radiative Transfer Framework for Spatially-Correlated Materials. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 37, 4 (July 2018), 83:1–83:13. <https://doi.org/10/gd52qg>
- James T. Kajiya. 1986. The Rendering Equation. *Computer Graphics (Proceedings of SIGGRAPH)* 20, 4 (Aug. 1986), 143–150. <https://doi.org/10/cvtf53j>
- Csaba Kelemen, László Szirmay-Kalos, György Antal, and Ferenc Csonka. 2002. A Simple and Robust Mutation Strategy for the Metropolis Light Transport Algorithm. *Computer Graphics Forum* 21, 3 (Sept. 2002), 531–540. <https://doi.org/10/bfrsqn>
- Jaroslav Krivánek, Iliyan Georgiev, Toshiya Hachisuka, Petr Vévoda, Martin Šík, Derek Nowrouzezahrai, and Wojciech Jarosz. 2014. Unifying Points, Beams, and Paths in Volumetric Light Transport Simulation. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 33, 4 (July 2014), 103:1–103:13. <https://doi.org/10/f6cz72>
- Peter Kutz, Ralf Habel, Yining Karl Li, and Jan Novák. 2017. Spectral and Decomposition Tracking for Rendering Heterogeneous Volumes. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 36, 4 (July 2017), 111:1–111:16. <https://doi.org/10/gbxjxg>
- Eric P. Lafortune and Yves D. Willems. 1993. Bi-Directional Path Tracing. In *Proceedings of the International Conference on Computational Graphics and Visualization Techniques (Compugraphics)*, H. P. Santo (Ed.), Vol. 93. Alvor, Portugal, 145–153.
- Jaakkko Leppänen. 2010. Performance of Woodcock Delta-Tracking in Lattice Physics Applications Using the Serpent Monte Carlo Reactor Physics Burnup Calculation Code. *Annals of Nuclear Energy* 37, 5 (May 2010), 715–722. <https://doi.org/10/d3fg8p>
- Michael D. McKay, Richard J. Beckman, and William Jay Conover. 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* 21, 2 (May 1979), 239–245. <https://doi.org/10/gfzncl>
- Bailey Miller, Iliyan Georgiev, and Wojciech Jarosz. 2019. A null-scattering path integral formulation of light transport. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 38, 4 (July 2019). <https://doi.org/10/gf6rz8>
- Ken Museth. 2013. VDB: High-Resolution Sparse Volumes with Dynamic Topology. *ACM Transactions on Graphics* 32, 3 (July 2013), 27:1–27:22. <https://doi.org/10/gfzq7s>
- Jan Novák, Iliyan Georgiev, Johannes Hanika, and Wojciech Jarosz. 2018. Monte Carlo Methods for Volumetric Light Transport Simulation. *Computer Graphics Forum (Proceedings of Eurographics State of the Art Reports)* 37, 2 (May 2018), 551–576. <https://doi.org/10/gd2jqq>
- Jan Novák, Andrew Selle, and Wojciech Jarosz. 2014. Residual Ratio Tracking for Estimating Attenuation in Participating Media. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 33, 6 (Nov. 2014), 179:1–179:11. <https://doi.org/10/f6r2nq>
- Mark Pauly, Thomas Kollig, and Alexander Keller. 2000. Metropolis Light Transport for Participating Media. In *Rendering Techniques (Proceedings of the Eurographics Workshop on Rendering)*, W. Hansmann, W. Purgathofer, François X. Sillion, Bernard Péroche, and Holly Rushmeier (Eds.). Springer-Verlag, Vienna, 11–22. <https://doi.org/10/gfzm93>
- Ken H. Perlin and Eric M. Hoffert. 1989. Hypertexture. *Computer Graphics (Proceedings of SIGGRAPH)* 23, 3 (July 1989), 253–262. <https://doi.org/10/fdmxsd>
- Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2016. *Physically Based Rendering: From Theory to Implementation* (3 ed.). Morgan Kaufmann, San Francisco, CA, USA.
- Hao Qin, Xin Sun, Qiming Hou, Baining Guo, and Kun Zhou. 2015. Unbiased Photon Gathering for Light Transport Simulation. *ACM Transactions on Graphics* 34, 6, Article 208 (Oct. 2015), 14 pages. <https://doi.org/10/f7wrc6>
- Matthias Raab, Daniel Seibert, and Alexander Keller. 2008. Unbiased Global Illumination with Participating Media. In *Monte Carlo and Quasi-Monte Carlo Methods*, Alexander Keller, Stefan Heinrich, and Harald Niederreiter (Eds.). Springer-Verlag, 591–605.
- Sheldon M. Ross. 1996. *Stochastic processes*. Wiley.
- Peter Shirley. 1991. Discrepancy as a Quality Measure for Sample Distributions. In *Proceedings of Eurographics*. Eurographics Association, Amsterdam, North-Holland, 183–194. <https://doi.org/10/gfznch>
- Jerome Spanier. 1966. Two Pairs of Families of Estimators for Transport Problems. *SIAM J. Appl. Math.* 14, 4 (July 1966), 702–713. <https://doi.org/10/dg35nt>

- Jerome Spanier and Ely Meyer Gelbard. 1969. *Monte Carlo Principles and Neutron Transport Problems*. Addison-Wesley, Reading.
- László Szirmay-Kalos, Balázs Tóth, and Milán Magdics. 2011. Free Path Sampling in High Resolution Inhomogeneous Participating Media. *Computer Graphics Forum* 30, 1 (March 2011), 85–97. <https://doi.org/10/bpkdzm>
- Eric Veach. 1997. *Robust Monte Carlo Methods for Light Transport Simulation*. Ph.D. Thesis. Stanford University, United States – California.
- Eric Veach and Leonidas J. Guibas. 1994. Bidirectional Estimators for Light Transport. In *Photorealistic Rendering Techniques (Proceedings of the Eurographics Workshop on Rendering)*, Georgios Sakas, Peter Shirley, and Stefan Müller (Eds.). Springer-Verlag, 145–167. <https://doi.org/10/gfznbh>
- Eric Veach and Leonidas J. Guibas. 1995. Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *Annual Conference Series (Proceedings of SIGGRAPH)*, Vol. 29. ACM Press, 419–428. <https://doi.org/10/d7b6n4>
- Eric Veach and Leonidas J. Guibas. 1997. Metropolis Light Transport. In *Annual Conference Series (Proceedings of SIGGRAPH)*, Vol. 31. ACM Press, 65–76. <https://doi.org/10/bkjgq4>
- Tien-Tsin Wong, Wah-Shing Luk, and Pheng-Ann Heng. 1997. Sampling with Hammersley and Halton Points. *Journal of Graphics Tools* 2, 2 (1997), 9–24. <https://doi.org/10/c936>
- E. R. Woodcock, T. Murphy, P. J. Hemmings, and T. C. Longworth. 1965. Techniques Used in the GEM Code for Monte Carlo Neutronics Calculations in Reactors and Other Systems of Complex Geometry. In *Applications of Computing Methods to Reactor Problems*. Argonne National Laboratory.

A EXPONENTIAL FORMULATION DERIVATION

Here we derive the classical exponential formulation of transmittance in Eq. (3). We begin by dividing both sides of Eq. (2) by $L(x)$:

$$-\frac{dL(x)}{L(x)dx} = -\frac{d \ln(L(x))}{dx} \stackrel{(2)}{=} -\mu(x). \quad (34)$$

Integrating from a to b gives (recall that $T(a, b) = L(a)/L(b)$)

$$-\left[\ln(L(x)) \right]_a^b = \ln\left(\frac{L(a)}{L(b)}\right) = \ln(T(a, b)) \stackrel{(34)}{=} - \int_a^b \mu(x) dx. \quad (35)$$

Finally, exponentiating the last equality in Eq. (35) yields Eq. (3).

B NULL-SCATTERING VOLTERRA DERIVATION

Here we derive the null-scattering extension (13) to our Volterra integral formulation (9). We start from Eq. (12), noting that it can be written as a canonical first-order linear differential equation:

$$-\frac{dL(x)}{dx} + \bar{\mu}(x)L(x) = \overline{\mu_n(x)L(x)}. \quad (36)$$

We now multiply both sides of Eq. (36) by the integrating factor $h(x) = \exp(-\int_a^x \bar{\mu}(x') dx')$, using its property $dh(x)/dx = -h(x)\bar{\mu}(x)$:

$$-h(x)\frac{dL(x)}{dx} - \frac{dh(x)}{dx}L(x) = -d\frac{h(x)L(x)}{dx} \stackrel{(36)}{=} h(x)g(x). \quad (37)$$

Integrating both sides of the last equality from a to b yields

$$-\int_a^b d\frac{h(x)L(x)}{dx} dx = h(a)L(a) - h(b)L(b) \stackrel{(37)}{=} \int_a^b h(x)g(x) dx. \quad (38)$$

Dividing by $h(a)L(b)$ and rearranging the terms yields an expression for $T(a, b) = L(a)/L(b)$:

$$T(a, b) = \frac{h(b)}{h(a)} + \int_a^b \frac{h(x)}{h(a)} \frac{g(x)}{L(b)} dx. \quad (39)$$

Substituting back $g(x)$ and $h(x)$, we finally obtain Eq. (13).

C VOLTERRA EXPANSION TO POWER SERIES

Here we derive our hypercube formulation (19) from our Volterra formulation (9). Expanding the recursion in Eq. (9) once gives

$$T(a, b) = 1 - \int_a^b \mu(x_1) dx_1 + \int_a^b \int_{x_1}^b \mu(x_1)\mu(x_2)T(x_2, b) dx_2 dx_1. \quad (40)$$

Continuing the expansion infinitely, we get

$$T(a, b) = 1 + \sum_{k=1}^{\infty} \int_a^b \int_{x_1}^b \cdots \int_{x_{k-1}}^b \left[\prod_{i=1}^k -\mu(x_i) \right] dx_k \cdots dx_2 dx_1 \quad (41a)$$

$$= 1 + \sum_{k=1}^{\infty} \int_{\mathcal{H}_k^s} \left[\prod_{i=1}^k -\mu(x_i) \right] dx = \sum_{k=0}^{\infty} \int_{\mathcal{H}_k^s} \left[\prod_{i=1}^k -\mu(x_i) \right] dx, \quad (41b)$$

where $dx = \prod_{i=1}^k dx_i$ is the standard k -dimensional differential volume measure. Each integral k in Eq. (41b) runs over a k -simplex

$$\mathcal{H}_k^s = \{ \mathbf{x} = (x_1, \dots, x_k) : a \leq x_1 \leq \cdots \leq x_k \leq b \} \subset \mathcal{H}_k \quad (42)$$

of a k -hypercube $\mathcal{H}_k = [a, b]^k$ with volume $|\mathcal{H}_k| = (b-a)^k$.

The coordinates x_i in every summand in Eq. (41) can be permuted freely without changing the value of the corresponding simplex integral, since the integrand is invariant to permutation. Each such permutation corresponds to integrating over one of $k!$ pair-wise congruent simplices with equal volumes $|\mathcal{H}_k|/k!$. We can extend each integral k in Eq. (41b) to run over the union of these $k!$ simplices, i.e. over the full k -hypercube \mathcal{H}_k , scaling the result by $1/k!$:

$$T(a, b) = \sum_{k=0}^{\infty} \int_{\mathcal{H}_k} \frac{1}{k!} \left[\prod_{i=1}^k -\mu(x_i) \right] dx = \sum_{k=0}^{\infty} \frac{1}{k!} \prod_{i=1}^k \left[- \int_a^b \mu(x) dx \right].$$

The LHS is identical to Eq. (19) and the RHS is identical to Eq. (15).

D EFFICIENCY OF SINGLE-TERM ESTIMATORS

We leverage our power-series formulation (26) to derive expressions for the variance and efficiency of the ratio-tracking, track-length, and residual-tracking estimators. We assume the upper and lower controls are proportional to the extinction, i.e. $\bar{\mu} \propto \underline{\mu} \propto \mu$, thus the null thickness τ_n is estimated with zero variance. The only source of variance are then the discrete sampling decisions that the estimators make.

Ratio-tracking estimator. The ratio-tracking estimator \widehat{T}_{rt} uses the pmf $P(k) = e^{-\bar{\tau}} \bar{\tau}^k / k!$ from Eq. (27) to choose a summand k in Eq. (22), and scores

$$\widehat{T}_{rt,k} = \frac{\tau_n^k}{\bar{\tau}^k}. \quad (43)$$

The variance of this estimator thus reads

$$V[\widehat{T}_{\text{rt}}] = \sum_{k=0}^{\infty} (\widehat{T}_{\text{rt},k})^2 \cdot P(k) - E^2[\widehat{T}_{\text{rt}}] = \sum_{k=0}^{\infty} \left(\frac{\tau_n^k}{\bar{\tau}^k} \right)^2 \cdot P(k) - e^{-2\tau} \quad (44a)$$

$$= \sum_{k=0}^{\infty} \frac{\tau_n^{2k}}{\bar{\tau}^{2k}} \cdot e^{-\bar{\tau}} \frac{\bar{\tau}^k}{k!} - e^{-2\tau} = e^{-\bar{\tau}} \sum_{k=0}^{\infty} \frac{\tau_n^{2k}}{\bar{\tau}^k} \frac{1}{k!} - e^{-2\tau} \quad (44b)$$

$$= e^{-\bar{\tau}} \sum_{k=0}^{\infty} \left(\frac{\tau_n^2}{\bar{\tau}} \right)^k \frac{1}{k!} - e^{-2\tau} = e^{-\bar{\tau}} e^{\frac{\tau_n^2}{\bar{\tau}}} - e^{-2\tau} \quad (44c)$$

$$= e^{\frac{\tau_n^2 - \bar{\tau}^2}{\bar{\tau}}} - e^{-2\tau} = e^{\frac{(\bar{\tau} - \tau_n)^2 - \bar{\tau}^2}{\bar{\tau}}} - e^{-2\tau} = e^{\frac{\tau^2 - 2\bar{\tau}\tau}{\bar{\tau}}} - e^{-2\tau} \quad (44d)$$

$$= e^{-2\tau + \frac{\tau^2}{\bar{\tau}}} - e^{-2\tau} = e^{-2\tau} \left(e^{\frac{\tau^2}{\bar{\tau}}} - 1 \right). \quad (44e)$$

The cost of every outcome $\widehat{T}_{\text{rt},k}$ in terms of number of medium extinction lookups is k , one lookup per null-thickness estimation sample x_i in Eq. (26). The expected cost $C[\widehat{T}_{\text{rt}}]$ is thus the mean of the used Poisson pmf, i.e. $\bar{\tau}$. The estimator's efficiency is then

$$\text{Eff}[\widehat{T}_{\text{rt}}] = \frac{1}{C[\widehat{T}_{\text{rt}}]V[\widehat{T}_{\text{rt}}]} = \frac{1}{\bar{\tau}e^{-2\tau} \left(e^{\frac{\tau^2}{\bar{\tau}}} - 1 \right)}. \quad (45)$$

Track-length estimator. The track-length estimator \widehat{T}_{tl} uses the same pmf (27) as ratio tracking. For a given index k , the estimator scores

$$\widehat{T}_{\text{tl},k} = \begin{cases} 1 & \text{with probability } P_k = \tau_n^k / \bar{\tau}^k, \\ 0 & \text{with probability } 1 - P_k. \end{cases} \quad (46)$$

With the assumption $\bar{\mu} \propto \mu$, the acceptance probability applied at each of the k integration samples x_i is $P = \mu_n(x_i)/\bar{\mu}(x_i) = \tau_n/\bar{\tau}$, and their product is P_k . The variance of \widehat{T}_{tl} reads

$$V[\widehat{T}_{\text{tl}}] = \sum_{k=0}^{\infty} (\widehat{T}_{\text{tl},k})^2 \cdot P(k) \cdot P_k - E^2[\widehat{T}_{\text{tl}}] \quad (47a)$$

$$= \sum_{k=0}^{\infty} 1 \cdot e^{-\bar{\tau}} \frac{\bar{\tau}^k}{k!} \cdot \frac{\tau_n^k}{\bar{\tau}^k} - e^{-2\tau} = e^{-\bar{\tau}} \sum_{k=0}^{\infty} \frac{\tau_n^k}{k!} - e^{-2\tau} \quad (47b)$$

$$= e^{-\bar{\tau}} e^{\tau_n} - e^{-2\tau} = e^{-\bar{\tau}} e^{\bar{\tau} - \tau} - e^{-2\tau} = e^{-\tau} - e^{-2\tau}. \quad (47c)$$

Interestingly, here the variance does *not* depend on the extinction upper control parameter.

The cost of each outcome $\widehat{T}_{\text{tl},k}$ is a random variable dependent on the acceptance probability $P = \tau_n/\bar{\tau}$ applied at each step $i = 1, \dots, k$. When the estimator terminates at the i th step, the cost is i . This happens when continuation is chosen at the previous $i-1$ steps, with probability P^{i-1} . Termination then occurs with probability $(1-P)$ when $i < k$ and with probability 1 when $i = k$. The expected cost of $\widehat{T}_{\text{tl},k}$ is the sum of cost times probability:

$$C[\widehat{T}_{\text{tl},k}] = \sum_{i=0}^{k-1} i \cdot P^{i-1} (1-P) + k \cdot P^{k-1}. \quad (48)$$

The pmf for sampling k is as in ratio tracking, $P(k) = e^{-\bar{\tau}} \bar{\tau}^k / k!$, and the expected cost of the estimator has a simple expression:

$$C[\widehat{T}_{\text{tl}}] = \sum_{k=0}^{\infty} C[\widehat{T}_{\text{tl},k}] \cdot e^{-\bar{\tau}} \frac{\bar{\tau}^k}{k!} = (1 - e^{-\tau}) \frac{\bar{\tau}}{\tau}. \quad (49)$$

Its efficiency is thus

$$\text{Eff}[\widehat{T}_{\text{tl}}] = \frac{1}{C[\widehat{T}_{\text{tl}}]V[\widehat{T}_{\text{tl}}]} = \frac{\tau}{\bar{\tau}e^{-3\tau}(e^\tau - 1)^2}. \quad (50)$$

Residual-tracking estimator. Residual tracking estimates Eq. (22) using a Poisson pmf $P(k)$ with mean $\bar{\tau} - \tau_n$ and thus scores

$$\widehat{T}_{\text{rrt},k} = e^{-\bar{\tau}} \frac{\tau_n}{\bar{\tau} - \tau_n}. \quad (51)$$

Its variance is derived analogously to that of ratio tracking:

$$V[\widehat{T}_{\text{rrt}}] = \sum_{k=0}^{\infty} (\widehat{T}_{\text{rrt},k})^2 \cdot P(k) - E^2[\widehat{T}_{\text{rrt}}] \quad (52a)$$

$$= \sum_{k=0}^{\infty} e^{-2\bar{\tau}} \frac{\tau_n^{2k}}{(\bar{\tau} - \tau_n)^{2k}} \cdot e^{-(\bar{\tau} - \tau_n)} \frac{(\bar{\tau} - \tau_n)^k}{k!} - e^{-2\tau} \quad (52b)$$

$$= e^{\frac{\tau^2 - 2\bar{\tau}\tau + \tau_n^2}{\bar{\tau} - \tau_n}} - e^{-2\tau}, \quad (52c)$$

where Eq. (52c) follows from simplifications similar to those in Eq. (44). The expected cost of the estimator is again equal its Poisson pmf mean, $\bar{\tau} - \tau_n$. Its efficiency thus reads

$$\text{Eff}[\widehat{T}_{\text{rrt}}] = \frac{1}{C[\widehat{T}_{\text{rrt}}]V[\widehat{T}_{\text{rrt}}]} = \frac{1}{(\bar{\tau} - \tau_n)(e^{\frac{\tau^2 - 2\bar{\tau}\tau + \tau_n^2}{\bar{\tau} - \tau_n}} - e^{-2\tau})}. \quad (53)$$

Note that when $\tau_n = 0$, Eq. (52c) reduces to Eq. (44e) and Eq. (53) reduces to Eq. (45).

Discussion. Equations (44) and (47) show that ratio tracking's variance is never higher than that of track-length's when $\bar{\tau} \geq \tau$, i.e. when $\bar{\mu}$ bounds the extinction. On the other hand, ratio tracking's cost of $\bar{\tau}$ is never lower than track-length's (49). Track-length can thus have higher efficiency, which is the case when $\bar{\mu}$ is relatively tight, i.e. in the upper regions of the plots in Fig. 7. Furthermore, for $0 \leq \tau_n \leq \tau \leq \bar{\tau}$ the variance of the residual-tracking estimator is never higher than that of ratio tracking and its efficiency is never lower.