

A Dual Theory of Inverse and Forward Light Transport

*Jiamin Bai
Manmohan Chandraker
Tian-Tsong Ng
Ravi Ramamoorthi*

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2010-101

<http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-101.html>

June 29, 2010



Copyright © 2010, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Acknowledgement

This work is funded by ONR YIP grant N00014-10-1-0032, ONR PECASE grant N00014-09-1-0741, a National Science Scholarship from A*STAR Graduate Academy of Singapore, as well as generous support from Adobe, NVIDIA, Intel and Pixar. We thank Joo Hwee Lim and Zhiyong Huang for kind support at the Institute for Infocomm Research and anonymous reviewers of the ECCV conference version for useful comments.

A Dual Theory of Inverse and Forward Light Transport

Jiamin Bai*

Manmohan Chandraker†

Tian-Tsong Ng‡

Ravi Ramamoorthi§

Abstract

A cornerstone of computer graphics is the solution of the rendering equation for interreflections, which allows the simulation of global illumination, given direct lighting or corresponding light source emissions. This paper lays the foundations for the inverse problem, whereby a dual theoretical framework is presented for inverting the rendering equation to undo interreflections in a real scene, thereby obtaining the direct lighting. Inverse light transport is of growing importance, enabling a variety of new applications like separation of individual bounces of the light transport, and projector radiometric compensation to display images free of global illumination artifacts in real-world environments that exhibit complex geometric and reflectance properties. However, solving the inverse problem involves the inversion of a large light transport matrix (acquired by measurement on real scenes). While straightforward matrix inversion is intractable for most realistic resolutions, there is scant prior work on either theoretical foundations or fast computational algorithms to meet the objectives of inverse light transport.

In this paper, we develop a mathematical theory that exposes the duality of forward and inverse light transport. Forward rendering also formally involves a matrix or operator inversion, which is conceptually equivalent to a multi-bounce Neumann series expansion. We show the existence of an analogous series for the inverse problem. However, the convergence is oscillatory in the inverse case, with more interesting conditions on material reflectance. Importantly, we give physical meaning to this duality, by showing that each term of our inverse series cancels an interreflection bounce, just as the forward series adds them.

In algorithmic terms, we develop the analog of iterative finite element methods like forward radiosity to efficiently solve light transport inversion. Our iterative inverse light transport algorithm is very fast, requiring only matrix-vector multiplications, and follows directly from the dual theoretical formulation. We also explore the connections to forward rendering in terms of Monte Carlo and wavelet-based techniques. As an initial practical application, we first acquire the light transport of a real static scene, and then demonstrate iterative inversion for radiometric compensation on high-resolution datasets, as well as rapid separation of the bounces of global illumination.

*e-mail: bjiamin@eecs.berkeley.edu

†e-mail: manukc@eecs.berkeley.edu

‡e-mail: ttng@i2r.a-star.edu.sg

§e-mail: ravir@eecs.berkeley.edu

1 Introduction

Forward global illumination, based on the theory of the rendering equation [Kajiya 1986], has been one of the great challenges and successes of computer graphics. Given the direct lighting (or light source emissions) in a virtual scene, interreflections and indirect light can efficiently be computed. In this paper, we consider the inverse problem [Seitz et al. 2005]—we seek to invert the rendering equation to undo the interreflections and recover the direct lighting in a real scene. For the forward problem, theoretical developments based on operator notation and error analysis, as well as algorithmic approaches such as efficient finite element radiosity and Monte Carlo methods are well-known [Arvo et al. 1994; Cohen and Wallace 1993; Kajiya 1986; Veach 1998]. However, relatively little is known about the theoretical and computational properties of the inverse problem. This paper develops a comprehensive theoretical analysis of inverse light transport, analyzes error and convergence, and demonstrates the inverse computational analogs to forward iterative finite element and Monte Carlo methods.

Motivation: We are motivated by two recent developments. First, many fast techniques for acquiring the light transport of real scenes have been proposed in recent literature [Debevec et al. 2000; Masselus et al. 2003; Peers et al. 2006]. Precomputed light transport is popular even for synthetic rendering [Sloan et al. 2002; Ng et al. 2003; Hasan et al. 2006]. In essence, these methods directly measure the effects of interreflections under various lighting conditions. The acquired light transport matrix has been used mainly for relighting applications, equivalent to matrix-vector multiplication.

However, another crucial problem is *inversion*, which enables new applications like illumination estimation, and separating direct lighting and each subsequent bounce of global illumination [Seitz et al. 2005]. Another important practical application is radiometric compensation,¹ when a projector is used to display an image on a scene with complex non-convex geometry and non-uniform reflectance, which may not be Lambertian. Interreflections are a serious issue in a real-world projector-camera system, and one would ideally like to compensate the projector image to remove global illumination effects (see Fig. 1).

Addressing these practical applications is conceptually simple—we invert the acquired light transport matrix for a real scene. However, the high resolution of real light

¹Another practical issue is to correct for geometric distortions, that we do not consider here, but calibrate for using standard techniques.

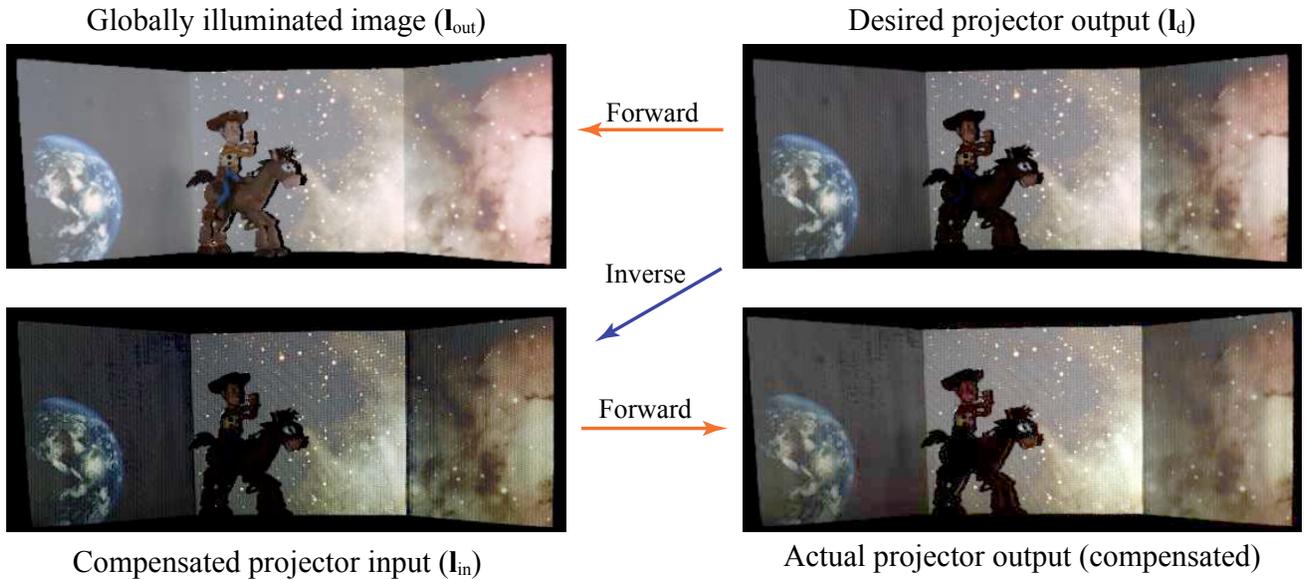


Figure 1: Application of inverse light transport for projector compensation in a real scene. This is suggestive of methods that could be used with projector-systems to achieve an artist’s desired appearance or look in interior scenes. **Top:** The desired projector output (right) leads to significant interreflections when displayed (left). **Bottom:** Our theory determines the pattern (left) whose projection is close to the desired (right). In effect, we have gone from global to local illumination, inverting or undoing interreflections. Our fast iterative method involves only matrix-vector multiplications, with each iteration taking only 0.03 sec. For a transport matrix of size $10^5 \times 10^5$, the full image \mathbf{l}_{in} is computed after several iterations in 2-3 secs.

transport data (from $10^3 \times 10^3$ in the simplest cases, up to $10^5 \times 10^5$ or higher) often makes standard matrix inversion impractical from both computation and memory standpoints. Most importantly, little is known about the theoretical properties or convergence of the inverse solution.

Contributions: We develop a novel theory for analyzing and solving inverse light transport. A perhaps surprising result is that there is a strong duality between forward and inverse rendering. This is because solving the (forward) rendering equation itself formally involves an operator or matrix inverse. Our main contribution is a theory that formulates forward and inverse light transport in similar ways, allowing us to leverage many theoretical results and algorithms from global illumination for the inverse problem.

Specifically, forward rendering readily admits to a Neumann series solution. In fact [Kajiya 1986] used this result to explain existing approximations and ray tracing. We develop a similar series expansion for inverse light transport, that explains recent work on fast stratified inversion [Ng et al. 2009]. Analyzing the convergence brings out subtle but important differences between these dual facets of light transport. Unlike in the forward case where physical bounces of light are added, the convergence of the inverse series is oscillatory. While the forward series convergence condition corresponds to energy conservation, in the inverse case the condition is more complex—a sufficient condition is that the albedo of surfaces is below 0.5, so that the net global illumi-

nation is still less than the direct lighting component.

An important theoretical contribution that highlights the duality is our derivation of a physical meaning for the inverse Neumann series. Just as each term of the forward Neumann series adds a bounce of light transport, we show that each term of the inverse series cancels a bounce.

The dual formulation leads to a new iterative inversion algorithm, analogous to the iterative finite element method for forward rendering, such as in radiosity. This is our main algorithmic or computational contribution. No formal matrix inversion is needed, and only matrix-vector multiplications are used. The resulting technique is thus very efficient, and can be related to standard Jacobi and Gauss-Seidel iterative methods. We also derive an iterative approach in case full matrix inversion is required. Finally, we explore Monte Carlo methods and final gather, showing the analogs between forward and inverse light transport.

We validate the theory and computational algorithms with numerical simulations and present a simple practical application as an example. Our practical method first requires the light transport of a static scene to be acquired. Our fast iterative inversion technique then enables full radiometric compensation of interreflections while projecting complex scenes (Fig. 1), as well as separation of individual local and global illumination components (Fig. 2).

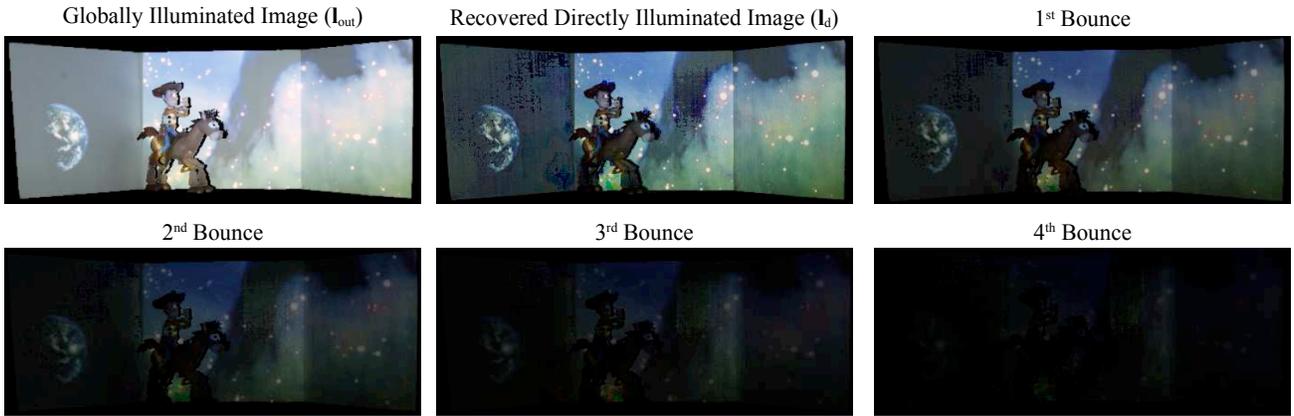


Figure 2: Separation of bounces of interreflection using our iterative light transport inversion technique. Each bounce is obtained in just 3 seconds for a $131K \times 131K$ light transport matrix.

2 Previous Work

Our work relates to light transport acquisition, the theory of forward and inverse rendering, computational methods for inversion, numerical linear algebra, and related practical applications. This paper is a more detailed version that supplements [Bai et al. 2010]. Most notably, we describe in detail the bounce cancellation of the inverse Neumann series in Section 5 and Monte Carlo algorithms in Section 8. In addition, we provide a fuller description of generality and limitations in Section 10.

Light Transport Acquisition: In this paper, we focus on the case where scene elements are illuminated individually by a single projector, with a single camera recording the output. This corresponds most closely to the setups in [Seitz et al. 2005; Peers et al. 2006; Ng et al. 2009]. It is somewhat distinct from relighting with distant illumination [Debevec et al. 2000], where a single lighting direction illuminates the entire surface. In fact, as we will discuss in Sec. 3, the difference is only in the first bounce or direct lighting, as interactions within the scene are still governed by the rendering equation. Extensions to incident (and reflected) light fields [Masselus et al. 2003; Sen et al. 2005; Garg et al. 2006] are encompassed by the theory, but not yet considered in our practical applications.

Inverse Rendering: Previous methods have considered inverting the direct reflection equation to acquire lighting and reflectance properties [Marschner 1998; Sato et al. 1999; Ramamoorthi and Hanrahan 2001]. [Yu et al. 1999] develop an inverse global illumination method for BRDF estimation. However, all these methods assume the scene geometry is known, and usually work with lower-resolutions for lighting, which makes analysis of interreflections much easier (and often requires only a few input images). In contrast, our work is closest to [Seitz et al. 2005], where only the light trans-

port matrix is observed—both geometry and reflectance are unknown, and are not explicitly estimated.

Forward Rendering: We draw on the rich history of global illumination by leveraging operator formulations and error analysis [Arvo et al. 1994], Monte Carlo algorithms [Veach 1998], and finite element radiosity methods [Cohen and Wallace 1993]. Many iterative radiosity techniques are also closely related to numerical linear algebra methods [Golub and van Loan 1996; Demmel 1997] for solving systems of linear equations, such as Jacobi and Gauss-Seidel iterations. Our framework enables similar relations to be drawn for inverse rendering. Similarly, our Monte Carlo method bears similarities to forward path tracing [Kajiya 1986], as well as von Neumann and Ulam’s original Monte Carlo matrix inversion method [Forsythe and Leibler 1950]. Future work could also consider analogs of hierarchical and wavelet radiosity [Hanrahan et al. 1991; Gortler et al. 1993] or photon mapping [Jensen 2001].

Computational Light Transport Inversion: Much of the most closely related work comes from radiometric compensation in projector-camera systems. [Wetzstein and Bimber 2007] form clusters of camera-projector pixels, doing a brute-force light transport inversion within clusters, but not considering inter-cluster interactions. This method is aimed at computational efficiency, but without clear error control. Iterative inverse methods for diffuse scenes are proposed in [Bimber et al. 2006]. More recently, [Ng et al. 2009] developed a series expansion for inverse light transport that they referred to as stratified inverses. We show that this series is a natural analog to the forward Neumann series. Our dual formulation enables us to go much further, clarifying the nature of convergence conditions. Most importantly, we derive a new computational analog to iterative finite element radiosity, as well as Monte Carlo methods. The resulting algorithms involve only matrix-vector (rather than matrix-

\mathbf{l}_{out}	Outgoing light (direct + interreflections)
\mathbf{l}_d	Direct light from sources or projector
\mathbf{l}_g	Global light from interreflections $\mathbf{l}_{\text{out}} = \mathbf{l}_d + \mathbf{l}_g$
\mathbf{l}_{in}	Incident lighting or projected pattern
\mathbf{S}	Operator/Matrix for forward transport
\mathbf{S}^{-1}	Operator/Matrix for inverse transport
\mathbf{R}	Operator/Matrix for interreflections only, $\mathbf{R} = \mathbf{S} - \mathbf{I}$
\mathbf{K}	Local reflection operator
\mathbf{G}	Geometric operator
\mathbf{A}	Net global transport, $\mathbf{A} = \mathbf{K}\mathbf{G}$
\mathbf{F}	First bounce from projector
\mathbf{T}	Observed light transport, $\mathbf{T} = \mathbf{S}\mathbf{F}$, $\mathbf{S} = \mathbf{T}\mathbf{F}^{-1}$
m	Norm of \mathbf{K} , related to maximum albedo ($m < 1$)
n	Number of terms in series expansion
k	Particular term in series or iteration
i_j	Index for point j on path
p	Probability for Monte Carlo sampling
f	Value of function (Monte Carlo expectation is f/p)
N	Transport resolution (matrix size is N^2)

Figure 3: Table of notation used in the paper.

matrix) multiplications, and are therefore significantly faster.

Practical Applications: Projector radiometric compensation has a long history [Nayar et al. 2003; Fujii et al. 2005; Ding et al. 2009], but these methods did not consider interreflections. This application also relates to techniques for making one object look like another [Raskar et al. 2001]. Our main practical contribution is a fast computation method for full light transport inversion.

Our other application is rapid direct and global separation for unstructured lighting. This relates most closely to [Nayar et al. 2006], who used a high-frequency illumination pattern and its complement. However, that method works only for a single image, where the entire scene is illuminated by the single light source or projector—not for the full light transport where the response for individual scene elements is computed, and where the incident illumination can come from many sources. On the other hand, we do first require acquisition of full light transport, unlike [Nayar et al. 2006]. Moreover, we can separate the different bounces of global illumination, like [Seitz et al. 2005], and can do so with much higher-resolution transport matrices at interactive rates.

3 Preliminaries

In this section, we lay out the problem statement, and discuss some practical issues for acquisition. A table of notation is given in Fig. 3.

Owing to the linearity of light transport and the rendering equation,

$$\mathbf{l}_{\text{out}} = \mathbf{S}\mathbf{l}_d, \quad (1)$$

where \mathbf{l}_{out} is the outgoing “global” light, and \mathbf{l}_d is the direct lighting on surfaces due to external sources. In continuous form, \mathbf{l}_{out} and \mathbf{l}_d are functions (of spatial location and outgoing direction), while \mathbf{S} is a linear operator that accounts for global illumination. If there are no interreflections, \mathbf{S} is the identity \mathbf{I} . When discretized for practical applications, \mathbf{l}_{out} and \mathbf{l}_d are vectors, while \mathbf{S} is the interreflection matrix. Equation 1 depends only on linearity, and holds for the light field, as well as a single camera view (image).

In traditional global illumination, \mathbf{l}_d is actually \mathbf{l}_e , the emission from sources. However, in our case, we do not see the light source or projector directly, but rather its effect on the scene or direct lighting (or equivalently “induced emission” \mathbf{l}_d). Equation 1 is also the same formulation used in recent direct-to-indirect transfer methods for relighting of synthetic scenes [Hasan et al. 2006].

The inverse light transport problem considered here is simply

$$\mathbf{l}_d = \mathbf{S}^{-1}\mathbf{l}_{\text{out}}, \quad (2)$$

where we seek to invert the operator or matrix \mathbf{S}^{-1} , undoing the effects of interreflections. Again, if there is no global illumination, $\mathbf{S} = \mathbf{S}^{-1} = \mathbf{I}$, and $\mathbf{l}_d = \mathbf{l}_{\text{out}}$.

Practical Issues: In practice, it is rare that \mathbf{S} is measured directly. Instead, a projector or illumination source lights the scene,

$$\mathbf{l}_d = \mathbf{F}\mathbf{l}_{\text{in}}, \quad (3)$$

where \mathbf{l}_{in} is the incident pattern projected (or distant light sources turned on), and \mathbf{F} is a “first-bounce” matrix or operator, that gives the direct lighting due to \mathbf{l}_{in} . We then observe

$$\mathbf{l}_{\text{out}} = \mathbf{T}\mathbf{l}_{\text{in}} = \mathbf{S}\mathbf{F}\mathbf{l}_{\text{in}}, \quad (4)$$

where the actual acquired light transport is $\mathbf{T} = \mathbf{S}\mathbf{F}$. The above expression holds for any light transport acquisition system, including projectors, distant and point light sources.

The remainder of the theoretical development in this paper focuses on analyzing and computing \mathbf{S}^{-1} . Eventual practical applications do need to convert from \mathbf{T} to \mathbf{S} , using

$$\mathbf{S} = \mathbf{T}\mathbf{F}^{-1}. \quad (5)$$

Moreover, applications like radiometric compensation actually seek to recover \mathbf{l}_{in} (rather than \mathbf{l}_d in equation 2) given by $\mathbf{l}_{\text{in}} = \mathbf{T}^{-1}\mathbf{l}_{\text{out}}$,

$$\mathbf{T}^{-1} = \mathbf{F}^{-1}\mathbf{S}^{-1} \quad \mathbf{l}_{\text{in}} = \mathbf{F}^{-1}\mathbf{l}_d. \quad (6)$$

Since we focus on global illumination \mathbf{S} , we will be interested in setups where \mathbf{S} is easy to obtain from \mathbf{T} , i.e., where \mathbf{F} is simple and at least approximately invertible.² Therefore,

²While the discussion in the paper is grounded in physical principles, from a numerical standpoint, \mathbf{F}^{-1} can also be seen simply as a preconditioner that improves the numerics of the matrix $\mathbf{S} = \mathbf{T}\mathbf{F}^{-1}$. Therefore, completely accurate estimation of \mathbf{F} is not required.

we consider projector-based acquisition, that illuminates a single spatial location and records the response, rather than light sources that illuminate the whole object (where \mathbf{F} is a low-pass filter, that is not easy to invert for diffuse surfaces [Ramamoorthi and Hanrahan 2001]). For projector-based acquisition, after geometric calibration, we can use the same parameterization for projection and camera images [Seitz et al. 2005]. \mathbf{F} is then a diagonal matrix, with \mathbf{F}^{-1} being trivial to compute, simply by taking reciprocals of the diagonal elements.³

At this stage, we note that \mathbf{F} need not correspond to the actual first bounce for an accurate light transport inversion. One may interpret light transport inversion in terms of general matrix inversion theory, whereby our choice of \mathbf{F} is simply Jacobi preconditioning, which is guaranteed to be convergent as long as \mathbf{T} is strictly diagonally dominant. See Section 10 for further discussion.

4 Dual Forward and Inverse Light Transport

At first glance, the inverse problem in equation 2 may be very different from forward light transport in equation 1. In this section, we show that the structure of the rendering equation exposes a strong duality between them. We then derive analogous Neumann series or expansions for forward and inverse transport. The next section briefly discusses convergence, followed by our main algorithmic contributions of fast inverse light transport algorithms in Sec. 7. Key theoretical results for each section are summarized in Fig. 4.

Using the operator form of the rendering equation [Arvo et al. 1994],

$$\mathbf{l}_{\text{out}} = \mathbf{l}_d + \mathbf{K}\mathbf{G}\mathbf{l}_{\text{out}}, \quad (7)$$

where \mathbf{K} considers the local reflection at a surface, governed by the BRDF, and \mathbf{G} is a geometric operator that transports outgoing to incident radiance. Note that this formulation is valid for any opaque BRDF when considering the full light field. While the theory we are about to present is fully general, our experiments will consider projection to a single view, which introduces practical limitations, as discussed in Sections 9 and 10. Denoting $\mathbf{A} = \mathbf{K}\mathbf{G}$, this can be written,

$$\mathbf{l}_{\text{out}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{l}_d, \quad (8)$$

from which it naturally follows that

$$\mathbf{S} = (\mathbf{I} - \mathbf{A})^{-1}. \quad (9)$$

This well known result shows that the forward problem formally involves a matrix or operator inversion, which indicates a similarity and duality with the inverse problem. Also

³In practice, we make the assumption that $\mathbf{F} = \text{diag}(\mathbf{T})$ similar to [Ng et al. 2009], i.e., using the diagonal elements of the transport after geometric calibration. This is an accurate approximation up to first order, since a surface point does not immediately interreflect onto itself.

note that if the scene geometry and reflectance (and hence \mathbf{A}) are known, we simply have $\mathbf{S}^{-1} = \mathbf{I} - \mathbf{A}$, as noted by [Seitz et al. 2005; Mukaigawa et al. 2006]. We focus here on cases where we only measure \mathbf{S} , but do not know or compute \mathbf{A} .

In fact, we can separate \mathbf{l}_{out} into the direct \mathbf{l}_d and indirect or global \mathbf{l}_g components,

$$\mathbf{l}_{\text{out}} = \mathbf{l}_d + \mathbf{l}_g = \mathbf{l}_d + \mathbf{R}\mathbf{l}_d, \quad (10)$$

which may be simplified to

$$\mathbf{l}_{\text{out}} = (\mathbf{I} + \mathbf{R})\mathbf{l}_d. \quad (11)$$

Here, we have defined another linear operator or matrix \mathbf{R} that accounts only for global illumination. By definition, \mathbf{R} is simply

$$\mathbf{R} = \mathbf{S} - \mathbf{I} \quad \mathbf{S} = \mathbf{I} + \mathbf{R}. \quad (12)$$

We are now ready to present an expression for inverse light transport, that is the dual to equation 9,

$$\mathbf{S}^{-1} = (\mathbf{I} + \mathbf{R})^{-1}. \quad (13)$$

The very similar or dual forms of equations 9 and 13 is a key insight in this paper, and allows direct leveraging of many forward rendering theories and algorithms for inverse rendering.

Neumann Forward and Inverse Series: It is well known that the forward equations 8 and 9 have series expansions corresponding physically to multiple bounces of light,

$$\mathbf{S} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \dots \quad (14)$$

We can also relate global illumination operator \mathbf{R} to this expansion,

$$\mathbf{R} = \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \dots = \mathbf{S} - \mathbf{I}. \quad (15)$$

Note that in our case, only \mathbf{S} (rather than \mathbf{A}) is known explicitly, and practical calculations simply use $\mathbf{R} = \mathbf{S} - \mathbf{I}$.

Mathematically, our dual formulation of inverse light transport in equation 13 has a series analogous to equation 14,

$$\mathbf{S}^{-1} = \mathbf{I} - \mathbf{R} + \mathbf{R}^2 - \mathbf{R}^3 + \dots \quad (16)$$

Note that the positive sign of \mathbf{R} implies the series is oscillatory. The physical meaning is harder to find than in the multi-bounce forward series. Intuitively, from equation 11, $\mathbf{l}_d = \mathbf{l}_{\text{out}} - \mathbf{R}\mathbf{l}_d$. But, evaluating the right hand side involves finding \mathbf{l}_d , which is unknown. So, we first approximate $\mathbf{l}_d \approx \mathbf{l}_{\text{out}}$, and remove all the global illumination due to $\mathbf{R}\mathbf{l}_{\text{out}}$, i.e. calculate $\mathbf{l}_d \approx \mathbf{l}_{\text{out}} - \mathbf{R}\mathbf{l}_{\text{out}}$. But, this overcompensates and gives too low a value, requiring higher-order

	Forward	Inverse
Problem	$\mathbf{l}_{\text{out}} = \mathbf{S}\mathbf{l}_{\text{d}}$	$\mathbf{l}_{\text{d}} = \mathbf{S}^{-1}\mathbf{l}_{\text{out}}$
Duality	$\mathbf{S} = (\mathbf{I} - \mathbf{A})^{-1}$	$\mathbf{S}^{-1} = (\mathbf{I} + \mathbf{R})^{-1}$
Series	$\mathbf{S} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots$	$\mathbf{S}^{-1} = \mathbf{I} - \mathbf{R} + \mathbf{R}^2 - \dots$
Bounces	$\mathbf{S}_n = \sum_{k=0}^n \mathbf{A}^k = \mathbf{S} + O(\mathbf{A}^{n+1})$	$\mathbf{S}_n^{-1} = \mathbf{I} - \mathbf{A} + O(\mathbf{A}^{n+1}) = \mathbf{S}^{-1} + O(\mathbf{A}^{n+1})$
Iteration	$\mathbf{l}_{\text{out}}^{(k)} = \mathbf{l}_{\text{d}} + \mathbf{A}\mathbf{l}_{\text{out}}^{(k-1)}$	$\mathbf{l}_{\text{d}}^{(k)} = \mathbf{l}_{\text{out}} - \mathbf{R}\mathbf{l}_{\text{d}}^{(k-1)}$
Monte Carlo	$\sum \mathbf{A}_{i_0 i_1} \mathbf{A}_{i_1 i_2} \dots \mathbf{l}_{\text{d}}(i_k)$	$\sum (-1)^k \mathbf{R}_{i_0 i_1} \mathbf{R}_{i_1 i_2} \dots \mathbf{l}_{\text{out}}(i_k)$

Figure 4: The duality of forward and inverse light transport, indicating analogous relations for some of the key properties. (Monte Carlo equations abbreviated; full forms in equations 46 and 47.)

corrections, and leading to the alternating signs in equation 16.

Finally, we note that equation 16 is (after suitable algebraic manipulations)⁴ identical to, and explains, the stratified inverses in [Ng et al. 2009]. Our derivation is simpler and directly relates to the rendering equation. This is much as the original rendering equation [Kajiya 1986] explained ray tracing as a special case of the forward series expansion. Carrying the analogy further, we will derive fast iterative algorithms (analogous to radiosity) in Sec. 7.

5 Inverse Neumann Series as Physical Bounces of Light

Forward light transport has an intuitive interpretation in terms of physical bounces of light, since each term of the forward Neumann series adds the next bounce. One may consider an approximation of order n :

$$\boxed{\mathbf{S}_n = \sum_{k=0}^n \mathbf{A}^k \quad \mathbf{S}_n - \mathbf{S} = O(\mathbf{A}^{n+1})} \quad (17)$$

In this section, we derive the interpretation of the inverse Neumann series in terms of physical bounces of light transport. A physical interpretation for the inverse series seems non-intuitive at first glance, since (16) is expressed in terms of \mathbf{R} , that includes all global illumination terms. Nevertheless, here we derive a surprising result: *each term of the inverse series cancels or zeros out the corresponding bounce of light transport, analogous to the forward case.*

We start with the basic relations, that

$$\mathbf{S} = (\mathbf{I} - \mathbf{A})^{-1}, \quad (18)$$

and that

$$\mathbf{S}^{-1} = (\mathbf{I} + \mathbf{R})^{-1}, \quad (19)$$

⁴In particular, note that $\mathbf{R} = \mathbf{S} - \mathbf{I}$, which is $\mathbf{T}\mathbf{F}^{-1} - \mathbf{I}$, or $\mathbf{T}\mathbf{T}^{(-1)} - \mathbf{I}$ in the notation of [Ng et al. 2009]. A final binomial expansion in $\mathbf{T}\mathbf{F}^{-1}$ or $\mathbf{T}\mathbf{T}^{(-1)}$, and using $\mathbf{T}^{-1} = \mathbf{F}^{-1}\mathbf{S}^{-1}$, enables one to derive their result.

where we also note that

$$\mathbf{R} = \mathbf{A} + \mathbf{A}^2 + \dots = \mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}. \quad (20)$$

Now, from equation 19 above, we can derive a series,

$$\mathbf{S}^{-1} = \sum_{k=0}^{\infty} (-1)^k \mathbf{R}^k = \sum_{k=0}^{\infty} (-1)^k [\mathbf{A}(\mathbf{I} - \mathbf{A})^{-1}]^k. \quad (21)$$

Note that we have used the final result of equation 20 in the last part. Moreover, while in general, raising a matrix (or operator) product to a power is complicated because of non-commutativity, in our case everything involves powers of \mathbf{A} , and so \mathbf{A} and $(\mathbf{I} - \mathbf{A})^{-1}$ commute, and can be exponentiated separately. Analogous to the order n approximation \mathbf{S}_n in the forward case, we can now write an expression for the corresponding approximation in the inverse case:

$$\mathbf{S}_n^{-1} = \sum_{k=0}^n (-1)^k \mathbf{R}^k = \sum_{k=0}^n (-1)^k \mathbf{A}^k (\mathbf{I} - \mathbf{A})^{-k}. \quad (22)$$

Binomial Series Expansion Using a standard binomial series expansion for $(\mathbf{I} - \mathbf{A})^{-k}$, this can be written as

$$\mathbf{S}_n^{-1} = \sum_{k=0}^n (-1)^k \mathbf{A}^k \sum_{l=0}^{\infty} \binom{k+l-1}{l} \mathbf{A}^l. \quad (23)$$

Our next step is to combine the powers of \mathbf{A} , using $m = l+k$ and $l = m - k$,

$$\mathbf{S}_n^{-1} = \sum_{k=0}^n \sum_{m=k}^{\infty} (-1)^k \binom{m-1}{m-k} \mathbf{A}^m. \quad (24)$$

It will simplify the later analysis if we treat $k = 0$ as a special case, given obviously from equation 22 as the identity. We also use $(m-1) - (m-k) = (k-1)$ in the combination,

$$\mathbf{S}_n^{-1} = \mathbf{I} + \sum_{k=1}^n \sum_{m=k}^{\infty} (-1)^k \binom{m-1}{k-1} \mathbf{A}^m. \quad (25)$$

To proceed further, we need to transpose the order of the summations. The outer summation should be about m ,

\mathbf{S}_n^{-1}	\mathbf{I}	\mathbf{A}	\mathbf{A}^2	\mathbf{A}^3	\mathbf{A}^4	\mathbf{A}^5	\mathbf{A}^6	\mathbf{A}^7
\mathbf{S}_0^{-1}	1	0	0	0	0	0	0	0
\mathbf{S}_1^{-1}	1	-1	-1	-1	-1	-1	-1	-1
\mathbf{S}_2^{-1}	1	-1	0	1	2	3	4	5
\mathbf{S}_3^{-1}	1	-1	0	0	-1	-3	-6	-10
\mathbf{S}_4^{-1}	1	-1	0	0	0	1	4	10
\mathbf{S}_5^{-1}	1	-1	0	0	0	0	-1	-5
\mathbf{S}_6^{-1}	1	-1	0	0	0	0	0	1
\mathbf{S}_7^{-1}	1	-1	0	0	0	0	0	0

$\mathbf{S}_n^{-1}\mathbf{S}$	\mathbf{I}	\mathbf{A}	\mathbf{A}^2	\mathbf{A}^3	\mathbf{A}^4	\mathbf{A}^5	\mathbf{A}^6	\mathbf{A}^7
$\mathbf{S}_0^{-1}\mathbf{S}$	1	1	1	1	1	1	1	1
$\mathbf{S}_1^{-1}\mathbf{S}$	1	0	-1	-2	-3	-4	-5	-6
$\mathbf{S}_2^{-1}\mathbf{S}$	1	0	0	1	3	6	10	15
$\mathbf{S}_3^{-1}\mathbf{S}$	1	0	0	0	-1	-4	-10	-20
$\mathbf{S}_4^{-1}\mathbf{S}$	1	0	0	0	0	1	5	15
$\mathbf{S}_5^{-1}\mathbf{S}$	1	0	0	0	0	0	-1	-6
$\mathbf{S}_6^{-1}\mathbf{S}$	1	0	0	0	0	0	0	1
$\mathbf{S}_7^{-1}\mathbf{S}$	1	0	0	0	0	0	0	0

Table 1: Coefficients of \mathbf{S}_n^{-1} and $\mathbf{S}_n^{-1}\mathbf{S}$. The series exhibit oscillatory convergence towards $\mathbf{I} - \mathbf{A}$ and \mathbf{I} respectively. The n term series is accurate up to \mathbf{A}^n , and in fact cancels or zeroes bounces up to that order, with errors only in higher-order terms or bounces $n + 1$ and higher.

which controls the powers. It is clear that we require $m \geq k$, which in turn leads to the relations that $k \leq m$ and (because we are consider the n term inverse series) that $k \leq n$,

$$\mathbf{S}_n^{-1} = \mathbf{I} + \sum_{m=1}^{\infty} \left[\sum_{k=1}^{\min(m,n)} (-1)^k \binom{m-1}{k-1} \right] \mathbf{A}^m. \quad (26)$$

Base Cases We treat the simple cases when $n = 0, 1$ and $m = 1$ first. When $n = 0$, the expression above just reduces to the identity (no bounce is cancelled as expected). When $n = 1$, only the $k = 1$ term is relevant, so we have

$$\mathbf{S}_1^{-1} = \mathbf{I} - \sum_{m=1}^{\infty} \mathbf{A}^m, \quad (27)$$

where we note that for $k = 1$, the $k - 1$ term in the combination reduces it to 1, and $(-1)^k = -1$. This is indeed the expected result, since $\mathbf{S}_1^{-1} = \mathbf{I} - \mathbf{R}$, and $\mathbf{R} = \mathbf{A} + \mathbf{A}^2 + \dots$

Finally, the special case $m = 1$ will be useful. In this case (assuming $n > 1$), the second summation in equation 26 will have upper limit $m = 1$, and the coefficient will simply be 1. Thus, for $n > 1$ (the cases $n = 0$ and $n = 1$ have already been dealt with),

$$\mathbf{S}_n^{-1} = \mathbf{I} - \mathbf{A} + \sum_{m=2}^{\infty} \left[\sum_{k=1}^{\min(m,n)} (-1)^k \binom{m-1}{k-1} \right] \mathbf{A}^m. \quad (28)$$

Zeroing of Higher-Order Bounces Now, consider the case when $m \leq n$. In this case, the second summation has a limit of $m > 1$, and the coefficient of \mathbf{A}^m becomes

$$\sum_{k=1}^m (-1)^k \binom{m-1}{k-1} = - \sum_{k'=0}^{m'} (-1)^{k'} \binom{m'}{k'} = 0. \quad (29)$$

where $m' = m - 1$ and $k' = k - 1$ (note this only works for $m > 1$; the $m = 1$ term is given as a special case in equation 28). The expression above is clearly 0, since those

are the coefficients in a binomial expansion of $(1 + x)^{m'}$, with $x = -1$.

This implies a key result, namely, that the \mathbf{A}^m terms vanish for $2 \leq m \leq n$, which in turn leads to

$$\boxed{\mathbf{S}_n^{-1} = \mathbf{I} - \mathbf{A} + O(\mathbf{A}^{n+1}) \quad \mathbf{S}_n^{-1} - \mathbf{S}^{-1} = O(\mathbf{A}^{n+1})} \quad (30)$$

where $O()$ denotes higher order terms, and $n > 1$. This means terms up to order n are correct, and in fact terms from $[\mathbf{A}^2 \dots \mathbf{A}^n]$ are 0.

Bounce Cancellation We have seen how higher-order terms are zeroed in the inverse operator series. We now show that applying the n -term inverse series to the original cancels the first n bounces. For this we write,

$$\mathbf{S}_n^{-1}\mathbf{S} = [(\mathbf{I} - \mathbf{A}) + O(\mathbf{A}^{n+1})] [\mathbf{I} - \mathbf{A}]^{-1}. \quad (31)$$

It is clear that the first part $\mathbf{I} - \mathbf{A}$ creates the identity as desired. The product $O(\mathbf{A}^{n+1})(\mathbf{I} - \mathbf{A})^{-1}$ is still of $O(\mathbf{A}^{n+1})$, since the inverse can be expanded in a Neumann series. Therefore,

$$\mathbf{S}_n^{-1}\mathbf{S} = \mathbf{I} + O(\mathbf{A}^{n+1}). \quad (32)$$

In other words, the n -term inverse series annihilates bounces $[1 \dots n]$, leaving only bounces $n + 1$ and higher.

Analytic Forms In fact, the inner summation in equation 28 can be performed symbolically (we did so using Mathematica), to derive

$$\mathbf{S}_n^{-1} = \mathbf{I} - \mathbf{A} + (-1)^n \sum_{m=n+1}^{\infty} \binom{m-2}{n-1} \mathbf{A}^m \quad (33)$$

$$\mathbf{S}_n^{-1}\mathbf{S} = \mathbf{I} + (-1)^n \sum_{m=n+1}^{\infty} \binom{m-1}{n} \mathbf{A}^m \quad (34)$$

Table 1 shows coefficients for $(m, n) \leq 7$. Figures 5 and 6 graphically illustrate the coefficients of \mathbf{S}_n^{-1} and $\mathbf{S}_n^{-1}\mathbf{S}$ respectively for $n \leq 12$ and $m \leq 7$. Owing to the $(-1)^n$ term, these coefficients oscillate until they are zeroed.

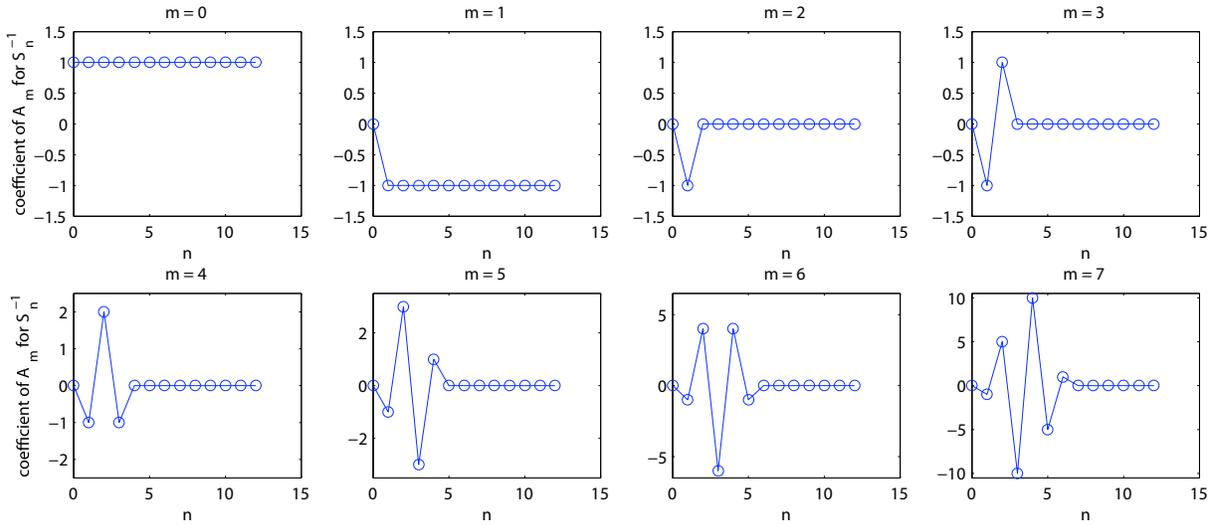


Figure 5: Coefficients of S_n^{-1} for $m = 0, \dots, 7$.

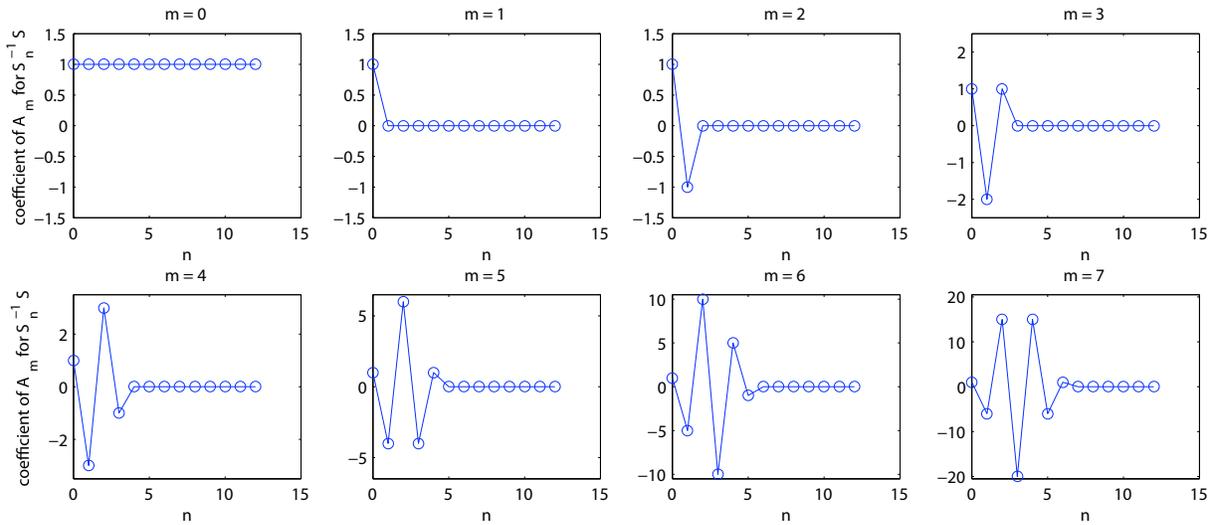


Figure 6: Coefficients of $S_n^{-1} S$ for $m = 0, \dots, 7$.

6 Convergence and Error Analysis

An immediate question is when the series in equation 16 converges, and how the error will decrease with more terms, as well as how this relates to the known properties of the forward series in equation 14.

For the forward case, [Arvo et al. 1994] prove several results, briefly summarized here. Since $\mathbf{A} = \mathbf{K}\mathbf{G}$, consider the norms of \mathbf{K} and \mathbf{G} first. $\|\mathbf{G}\| = 1$ for a closed enclosure (less for open scenes), while $\|\mathbf{K}\| = m < 1$, where m relates to the albedo of the surface (for non-diffuse materials, it is the maximum over all incident directions of the fraction of total energy reflected). The relation $m < 1$ comes from

energy conservation, excluding perfect reflectors,⁵

$$\|\mathbf{K}\| \leq m < 1 \quad \|\mathbf{A}\| \leq m < 1. \quad (35)$$

The last relation follows because $\|\mathbf{A}\| \leq \|\mathbf{K}\| \|\mathbf{G}\|$. In discrete terms, the matrix $\mathbf{I} - \mathbf{A}$ is diagonally dominant. Since $\|\mathbf{A}\| < 1$, the forward series always converges for physical environments.

We now turn to the inverse series in equation 16, and we desire that $\|\mathbf{R}\| < 1$ for convergence. A bound from equation 15 is,

$$\|\mathbf{R}\| \leq \|\mathbf{A}\| + \|\mathbf{A}^2\| + \dots \leq m + m^2 + \dots = \frac{m}{1 - m}. \quad (36)$$

⁵Note that these relations hold in any L^p norm, since by reciprocity, $\|\mathbf{K}\|_1 = \|\mathbf{K}\|_\infty = m$, and $\|\cdot\|_p \leq \max(\|\cdot\|_1, \|\cdot\|_\infty)$.

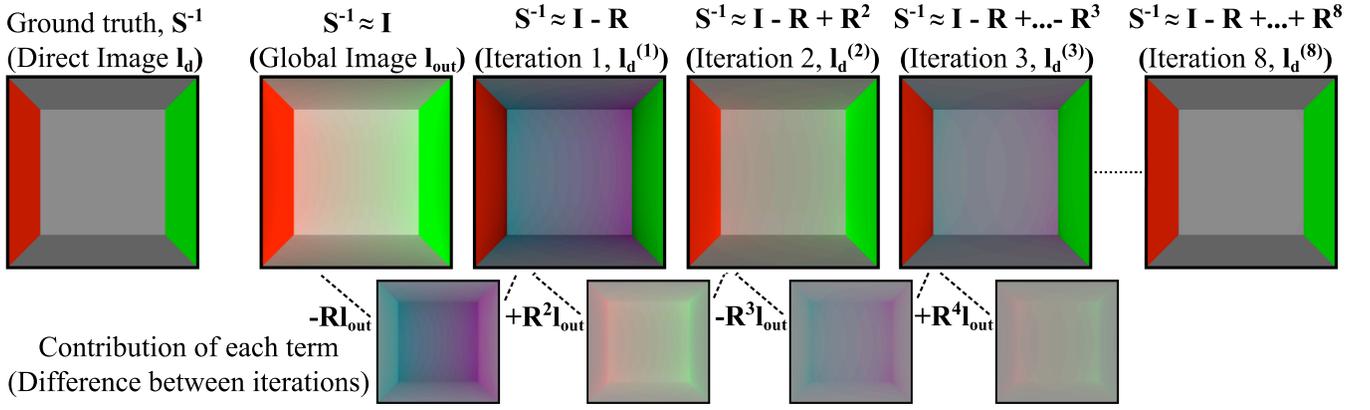


Figure 7: Top: From left to right, we add more terms of the inverse series, going from the simulated global illumination output I_{out} to the “direct lighting” result I_d (shown leftmost). These terms also correspond to the iterations introduced in Sec. 7. The results oscillate between over and under-compensation, converging after 4-8 iterations. **Bottom:** Contributions of individual terms (neutral grey is 0). Odd iterations over-compensate interreflections, giving rise to cyan and magenta colors, while even iterations under-compensate retaining some red and green color bleeding. As can be seen, the corrections of successive terms (iterations) reduce rapidly, converging to direct lighting only.

Setting $\frac{m}{1-m} < 1$, we obtain,

$$\|\mathbf{R}\| < 1 \quad \text{if} \quad m < \frac{1}{2}. \quad (37)$$

Intuitively, if the diffuse albedo (or maximum fraction of energy reflected for any incident direction for non-diffuse materials) is less than $1/2$, the norm of the total global illumination operator $\|\mathbf{R}\|$ is still less than that of the direct lighting operator \mathbf{I} . In matrix terms, $\mathbf{S} = \mathbf{I} + \mathbf{R}$ is diagonally dominant. Because the inverse series is oscillatory, we require to bound the full global illumination, rather than just each bounce as in the forward case. Note however, that equation 37 is only a sufficient, but not a necessary condition.

Error Analysis: The series expansions are usually truncated to a finite number n of terms (i.e. approximate \mathbf{S} as \mathbf{S}_n). The error introduced by doing so can easily be bounded. In the forward case,

$$\begin{aligned} \|\mathbf{S} - \mathbf{S}_n\| &\leq \sum_{k=n+1}^{\infty} \|\mathbf{A}^k\| \leq \sum_{k=n+1}^{\infty} m^k \\ &= \frac{m^{n+1}}{1-m}. \end{aligned} \quad (38)$$

Similarly, for the inverse series,

$$\begin{aligned} \|\mathbf{S}^{-1} - \mathbf{S}_n^{-1}\| &\leq \sum_{k=n+1}^{\infty} \|\mathbf{R}^k\| \leq \sum_{k=n+1}^{\infty} \left(\frac{m}{1-m}\right)^k \\ &= \frac{m^{n+1}}{(1-m)^n(1-2m)}. \end{aligned} \quad (39)$$

Numerical Simulations: We verify our results and provide insights with numerical simulations. For simplicity, we

consider a diffuse box, without shadows but with interreflections. This is a closed environment, so that $\|\mathbf{G}\| = 1$. For greatest accuracy, \mathbf{A} (and hence \mathbf{S}) is computed directly using analytic form factors for a discretized mesh [Schröder and Hanrahan 1993].

Figure 7 assumes that I_d is constant on each surface (but with different albedos and red-green walls as in a conventional Cornell Box). I_{out} includes global illumination effects. From left to right, we see the addition of more terms from equation 16 that oscillate between over and under-compensating the interreflections, till we converge to I_d . Interestingly, while forward global illumination in I_{out} results in predictable red and green color-bleeding, odd terms (or iterations in Sec. 7) of the inverse series over-compensate interreflections and give rise to cyan and magenta colors instead. The final inverse light transport solution for I_d has no color bleeding as desired.

In Fig. 8, we analyze errors and convergence. Fig. 8a shows error at a single point, that clearly indicates oscillatory convergence. While there is somewhat more error near corners, the corners, edges and face centers converge in largely similar ways. Fig. 8b compares error for the whole \mathbf{S}^{-1} operator with the theoretical bound in equation 39. Excellent agreement in form is obtained, with the bound being loose only by a constant factor. Because convergence is exponential (as per a geometric series), the vertical axis is on a logarithmic scale for this and following graphs. Next, Fig. 8c considers the effects of different albedos. Convergence rate varies inversely with albedo, as predicted by theory. Even for albedos like 0.45, close to the theoretical limit, at most 10-20 terms or iterations suffices. However, albedos very close to 0.5 show very slow convergence, and those greater than 0.51 diverge. Note that this figure is for a fully closed box. Fig-

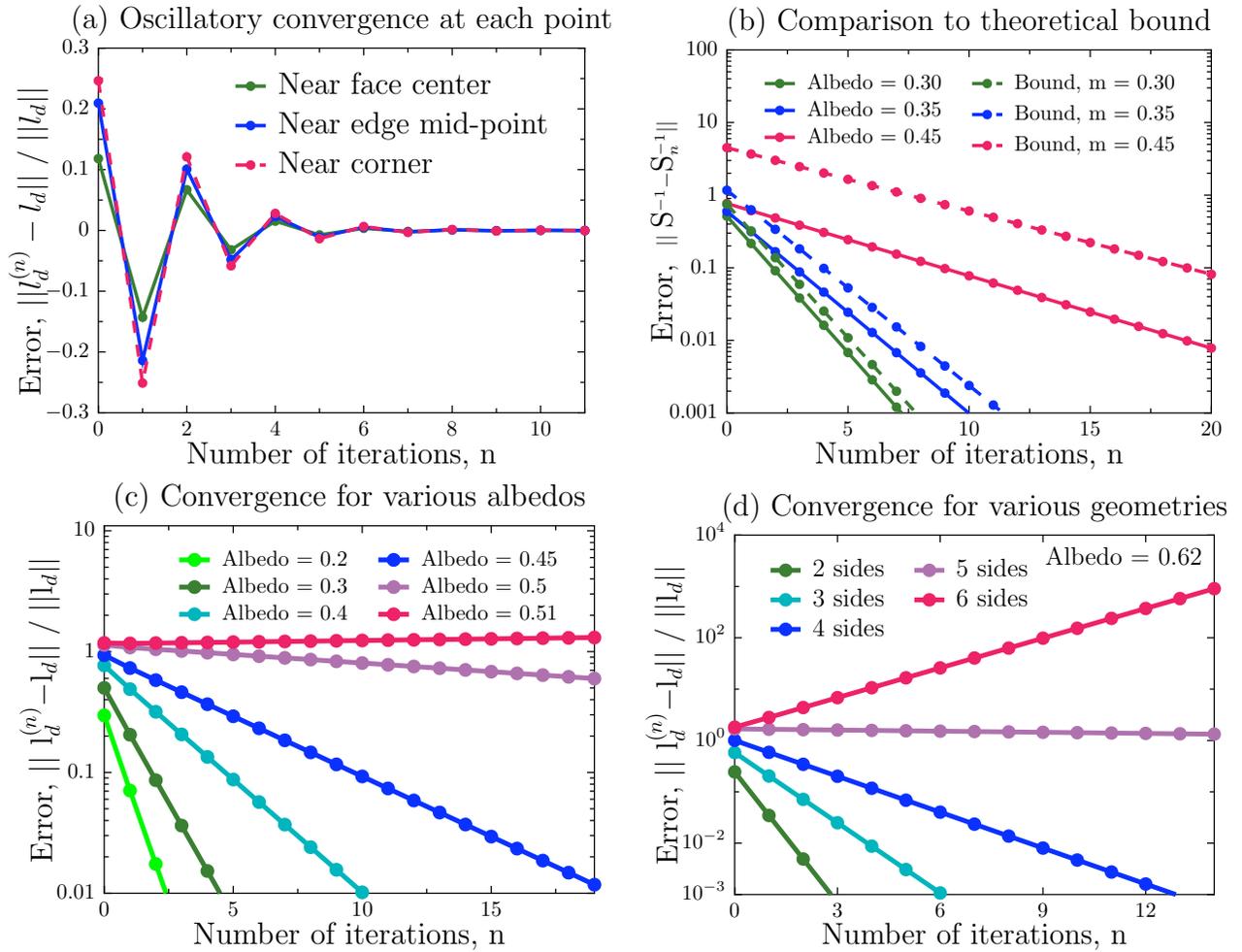


Figure 8: Error analysis for convergence of inverse series. **(a):** Convergence at different points (center, edge, corner), showing largely similar behavior. **(b):** Comparison of error to theoretical bound for different albedos showing good agreement. **(c):** Convergence for different albedos. As predicted by theory, convergence is faster for lower albedos, upto the limit of 0.5. An albedo of 0.51 leads to divergence. **(d):** An albedo of 0.62 diverges for a closed box (6 sides) and shows very slow convergence as expected for a box with 5 sides, but rapid convergence for more open environments (fewer sides, smaller $\|\mathbf{G}\|$).

ure 8d shows more realistic cases of partially open environments (fewer than 6 sides for the box), where $\|\mathbf{G}\| < 1$. For an albedo of 0.62, close to the theoretical limit for a 5-sided box, very slow convergence is achieved for 5 sides (and divergence for 6) as expected, but rapid convergence for more open environments.

Finally, Fig. 9 shows a scene with occlusions and glossy surfaces. Similar behaviors hold as above, with convergence of the inverse series to direct lighting even for high gloss where the $m < 0.5$ condition is not strictly satisfied.

7 Fast Iterative Computation

In this section, we introduce our main algorithmic contribution—a fast iterative method to compute inverse light transport, using only matrix-vector multiplications.

This method is the dual to iterative forward rendering methods like finite element radiosity, and we also explore analogous wavelet accelerations.

For forward rendering, one rarely computes the series in equation 14 to explicitly determine \mathbf{S} . This is mainly because of the high cost of matrix-matrix multiplications on high-resolution scenes. Instead, finite element and radiosity methods [Cohen and Wallace 1993] try to solve

$$\mathbf{l}_{\text{out}} = \mathbf{l}_d + \mathbf{A}\mathbf{l}_{\text{out}} \quad (40)$$

iteratively, which corresponds directly to equation 7. This iteration is numerically stable, and requires only the matrix-vector multiplication for $\mathbf{A}\mathbf{l}_{\text{out}}$. Each step computes

$$\mathbf{l}_{\text{out}}^{(k)} = \mathbf{l}_d + \mathbf{A}\mathbf{l}_{\text{out}}^{(k-1)}, \quad (41)$$

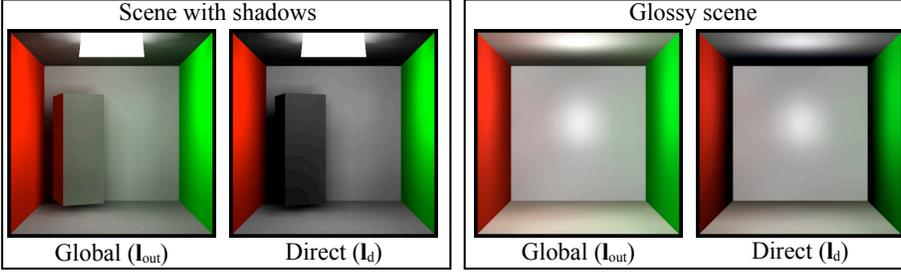


Figure 9: Validation of the theory for shadowed and non-Lambertian scenes. The iterative method of Section 7 recovers \mathbf{I}_d in 10 iterations for the shadowed scene and 20 for the glossy one.

where the superscript stands for the step k , and $\mathbf{I}_{\text{out}}^{(0)} = \mathbf{I}_d$. It is important to note that n steps correspond simply to computing the effect of the first n terms of the series in equation 14.

For inverse rendering, one can derive a similar relation,

$$\mathbf{I}_d = \mathbf{I}_{\text{out}} - \mathbf{R}\mathbf{I}_d, \quad (42)$$

which is analogous to equation 40, and follows from equation 11. The iterative solution naturally follows, dual to equation 41,

$$\mathbf{I}_d^{(k)} = \mathbf{I}_{\text{out}} - \mathbf{R}\mathbf{I}_d^{(k-1)}, \quad (43)$$

where again the first n steps correspond to the first n terms in equation 16. Note the negative sign on \mathbf{R} (compare to $+\mathbf{A}$ in equation 41), corresponding to the oscillatory nature of the series.

As written, equation 43 (and equation 41) corresponds to the Jacobi iteration for solving systems of linear equations [Golub and van Loan 1996]. If the \mathbf{I}_d are updated in place (instead of at the end of a step), this is the Gauss-Seidel method. Both techniques are popular in forward radiosity. By framing inverse light transport as dual to forward transport, we could also leverage other computational methods in future, such as Southwell iteration, successive over-relaxation, and conjugate gradient solutions.

Matrix Iteration: While rarely used in forward rendering, there is also a corresponding iteration for the full matrix or operator, in cases where we seek to precompute \mathbf{S} or \mathbf{S}^{-1} . In the forward case,

$$\mathbf{S}_k = \mathbf{I} + \mathbf{A}\mathbf{S}_{k-1}, \quad (44)$$

with $\mathbf{S}_0 = \mathbf{I}$. Correspondingly, in the inverse case,

$$\mathbf{S}_k^{-1} = \mathbf{I} - \mathbf{R}\mathbf{S}_{k-1}^{-1}. \quad (45)$$

These iterations are not significantly more efficient than a proper factorization for computing equations 14 and 16 directly, but do provide an elegant and numerically stable iterative scheme.

Wavelet and Hierarchical Methods: The matrix-vector multiplication $\mathbf{R}\mathbf{I}_d$ in equation 43 is the time-consuming step. We can wavelet-transform and approximate the vector \mathbf{I}_d , as well as the rows of \mathbf{R} , to speed up the matrix-vector multiply. This is analogous to wavelet radiosity and light transport in forward rendering [Gortler et al. 1993; Ng et al. 2003]. Other hierarchical approaches, analogous to [Hanrahan et al. 1991], can also be explored.

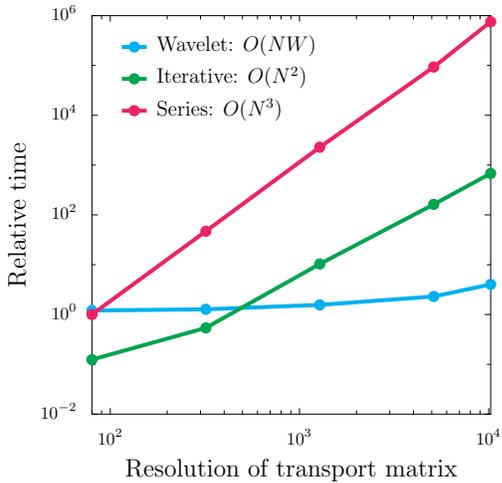
Numerical Simulations: As the baseline, we use matrix-matrix multiplications to directly compute the series in equation 16 (explicit matrix inversion is intractable for high resolutions). In Fig. 10, we compare to the standard iteration in equation 43, and accelerations using wavelets. The series method scales as $O(N^3)$, where N is the transport resolution and rapidly becomes impractical. The iterative method uses only matrix-vector multiplications and is much faster $O(N^2)$, with a speedup of three orders of magnitude for large sizes. Wavelet acceleration theoretically leads to linear $O(NW)$ performance, where the number of wavelets W in each row is relatively insensitive to N . The benefits are more noticeable at higher resolutions, where wavelet sparsity W outweighs the transform overhead—wavelets would provide significant savings at the resolutions in many real experiments.

8 Monte Carlo Algorithms

Besides finite element methods like radiosity, forward rendering has developed a suite of Monte Carlo techniques. In fact, [Kajiya 1986] proposed that the forward Neumann series in equation 14 could be solved with Markov Chain Monte Carlo or path tracing. Treating \mathbf{A} as a matrix, we need to consider all permutations of indices,

$$\mathbf{I}_{\text{out}}(i_0) = \mathbf{I}_d(i_0) + \sum_{k=1}^{\infty} \sum_{i_1, i_2, \dots, i_k} \mathbf{A}_{i_0 i_1} \mathbf{A}_{i_1 i_2} \dots \mathbf{A}_{i_{k-1} i_k} \mathbf{I}_d(i_k), \quad (46)$$

where the first summation is over all terms k in the series, or all path lengths in a path tracing context. The different indices correspond to all matrix sums, or equivalently all paths, where each i_j chooses a particular point on the path. In Monte Carlo path tracing, essentially the above form is implicitly used, but the \mathbf{A} matrix is not usually computed



Method	Resolution of transport matrix (N)				
	80	320	1280	5120	10240
Series	1.0	47.0	2.3e3	9.3e4	7.5e5
Iterative	0.1	0.5	10.3	162.5	679.0
Wavelet	1.2	1.3	1.6	2.3	4.0

Normalization: 1.0 = 5.57e-4 seconds

Figure 10: Timings for series, iterative finite element, and wavelet accelerated methods (using Daubechies4 wavelets). N is the transport resolution (matrix is of size N^2). We normalize timings so that 1.0 corresponds to 5.57×10^{-4} seconds, with experiments in Matlab on an Intel i7 machine. All methods were run to an error of 1%.

explicitly, and elements of it are generated on the fly.

The inverse series in equation 16 has an exactly analogous form,⁶

$$\mathbf{I}_d(i_0) = \mathbf{I}_{\text{out}}(i_0) + \sum_{k=1}^{\infty} (-1)^k \sum_{i_1, i_2, \dots, i_k} \mathbf{R}_{i_0 i_1} \mathbf{R}_{i_1 i_2} \dots \mathbf{R}_{i_{k-1} i_k} \mathbf{I}_{\text{out}}(i_k), \quad (47)$$

where the oscillatory behavior requires the additional $(-1)^k$ factor. A direct Monte Carlo algorithm is to use a number of samples, for each of which the indices i_1, i_2, \dots, i_k are drawn at random. The expectation of these samples then gives the desired result. Our implementation makes a number of optimizations, that correspond to analogous techniques in forward rendering.

First, for each sample, we choose a path length k . We assign probabilities to different path lengths in proportion to their expected contribution, which decays with k . From the convergence and error analysis in equation 39, we use the normalized probability

$$p(k) = \left(\frac{m}{1-m} \right)^k \frac{(1-2m)}{m}, \quad (48)$$

⁶Equations 46 and 47 are abbreviated for brevity in the table in Fig. 4.

with m being an estimate of the average albedo of the scene. We next choose indices i_1, i_2, \dots, i_k . These can be chosen randomly, or we can use importance sampling on each row of the matrix \mathbf{R} ,

$$p(i_j | i_{j-1}) = \frac{\mathbf{R}_{i_{j-1} i_j}}{\sum_{i_p} \mathbf{R}_{i_{j-1} i_p}} = \frac{\mathbf{R}_{i_{j-1} i_j}}{|\mathbf{R}_{i_{j-1}}|}, \quad (49)$$

where we normalize by the sum of elements over the full row, and the last step simply denotes the row sum more compactly as $|\mathbf{R}_{i_{j-1}}| = \sum_{i_p} \mathbf{R}_{i_{j-1} i_p}$. These row sums correspond to a generalized analog of albedos or BRDFs along the path, suitably weighted. Note that importance sampling in forward rendering is usually based on some partial information like lighting or BRDF, but we have the luxury of the full accurate \mathbf{R} matrix to importance sample.⁷ This greatly simplifies the final expressions.

Finally, the net image is just the expected value over all paths/samples. For each path, we must divide the value f by the probability (in this case, f is simply the appropriate term in the summation on the right-hand side of equation 47). Since f involves expressions of the form $\mathbf{R}_{i_{j-1} i_j}$, they cancel with the probabilities above as they should for good importance sampling,

$$\frac{f(k; i_1, i_2, \dots, i_k)}{p(k; i_1, i_2, \dots, i_k)} = \frac{(-1)^k}{p(k)} |\mathbf{R}_{i_0}| |\mathbf{R}_{i_1}| \dots |\mathbf{R}_{i_{k-1}}| \mathbf{I}_{\text{out}}(i_k), \quad (50)$$

where we must average (take the expected value) over all samples to obtain \mathbf{I}_d , and also add the initial term $\mathbf{I}_{\text{out}}(i_0)$ per equation 47. Note that the above simplified form is valid only if we importance sample properly when choosing the next index along a path.

Hybrid Methods: Besides the above pure Monte Carlo path tracing analog, we can also explore hybrids of iterative and Monte Carlo techniques. Note that these types of hybrids are rarely used in forward rendering but follow naturally from our framework. For example, we can speed up the matrix-vector multiplication $\mathbf{R}\mathbf{I}_d$ in equation 43 by Monte Carlo sampling only some of the columns for each row, using the importance sampling scheme above. We can also explore an analogy to final gather in forward rendering, where we use fewer samples for the iteration, but then compute the final step with a direct matrix-vector multiplication.

Numerical Simulations: Figure 11 first demonstrates that variance varies inversely with the number of samples per pixel as expected (each “pixel” corresponds to the intensity of an area element on the box in Fig. 7). The images in the top row show the power of final gather—Monte Carlo with 30 samples is noisy as expected, but is smoothed out almost

⁷Note that building the probability tables for importance sampling does require a preprocess for each row of the matrix. This preprocess is done once, after acquisition and before any specific \mathbf{I}_{out} is chosen.

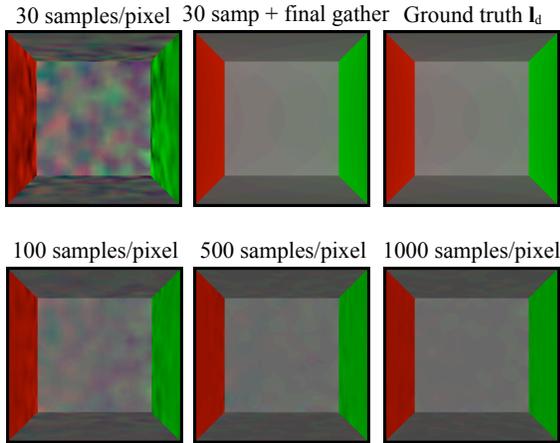
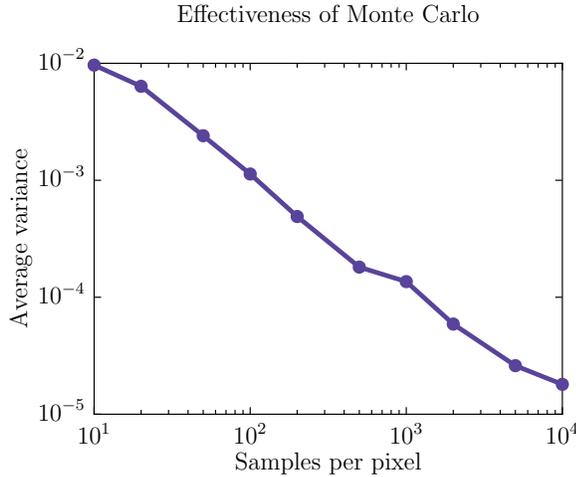


Figure 11: Top: Graph of variance in Monte Carlo methods, which shows the expected behavior, varying inversely with the number of samples per pixel. **Bottom:** In the top row, we show that only 30 samples per pixel (that in itself is extremely noisy) is adequate to produce good results using final gather. In the bottom row, as expected, Monte Carlo becomes more accurate with more samples. The transport resolution N in these cases is 5120.

completely using one direct iteration (the final gather). In the bottom row, we see that as expected, pure Monte Carlo converges as the number of samples is increased.⁸

9 Experiments with Real Data

In this section, we illustrate the application our iterative inverse light transport algorithm for two scenarios—separating the bounces of light transport and projector radiometric compensation. The accuracy of our algorithms is established by a few didactic examples, while their computational utility is

⁸Our current Matlab implementation is not optimized for the sampling process, making a direct timing comparison to finite elements difficult. Hence, we simply report on number of samples per pixel.

demonstrated by performance on high resolution transport matrices. See Section 10 for a discussion of limitations imposed by our choice of experimental conditions.

Acquisition Details: Our acquisition setup consists of a Dell 4310WX projector and a Canon EOS 5D Mark II camera. An accurate, one-time, radiometric calibration of the projector and camera response curves is performed to ensure linearity of the corresponding signals. While prior work has obtained transport matrices at resolutions comparable to ours [Sen et al. 2005], it has mainly been for applications akin to relighting. In contrast, the inverse problems that form our application domain require greater fidelity in the elements of the transport matrix. Thus, a judicious consideration of signal to noise ratio is necessary to capture as many of the weaker interreflection bounces as possible while discarding the sensor noise. To faithfully capture the energy of the transport matrix, up to 8 images at various exposures are assembled into a high dynamic range image. The projector’s black offset is computed at the highest exposure to average out high frequency fluctuations. For the higher resolution scenes, a hierarchical subdivision scheme is used to simultaneously acquire portions of the transport matrix which are not in mutual conflict, similar to [Sen et al. 2005].

Projector Radiometric Compensation: The ubiquitous use of projectors may necessitate inverting photometric distortions and interreflection effects to simulate any desired appearance in non-flat, non-Lambertian spaces. In terms of our theory, given a desired appearance \mathbf{I}_{out} , we seek to invert the light transport to find $\mathbf{I}_{\text{d}} = \mathbf{S}^{-1}\mathbf{I}_{\text{out}}$. As discussed in Sec. 3, we must account for the first bounce \mathbf{F} from the projector, and actually compute $\mathbf{I}_{\text{in}} = \mathbf{T}^{-1}\mathbf{I}_{\text{out}}$.

Fig. 1 shows results for radiometric compensation to project a desired image onto a scene with non-Lambertian materials, occlusions and interreflections. Clearly, the desired appearance is closely matched. The size of the transport matrix is $131K \times 131K$, for which our iterative algorithm performs radiometric compensation in only about 3 secs. While such high resolutions may be infeasible for a straightforward matrix inversion, based on the patterns in Fig. 10, the stratified inverses method of [Ng et al. 2009] will require 1 – 2 orders of magnitude more time. Also, in contrast to the method of [Wetzstein and Bimber 2007], our algorithms are physically motivated and not contingent on any tunable parameters.

Separating Bounces: One consequence of our theory is that once the light transport has been acquired, we can quickly separate an image into the different bounces (direct, 1st bounce indirect, 2nd bounce indirect and so on). It follows from (41), noting that $\mathbf{S}^{-1} = \mathbf{I} - \mathbf{A}$, that the k -th indirect bounce is

$$\mathbf{I}_{\text{out}}^{(k+1)} - \mathbf{I}_{\text{out}}^{(k)} = \mathbf{I}_{\text{d}} - \mathbf{S}^{-1}\mathbf{I}_{\text{out}}^{(k)}. \quad (51)$$

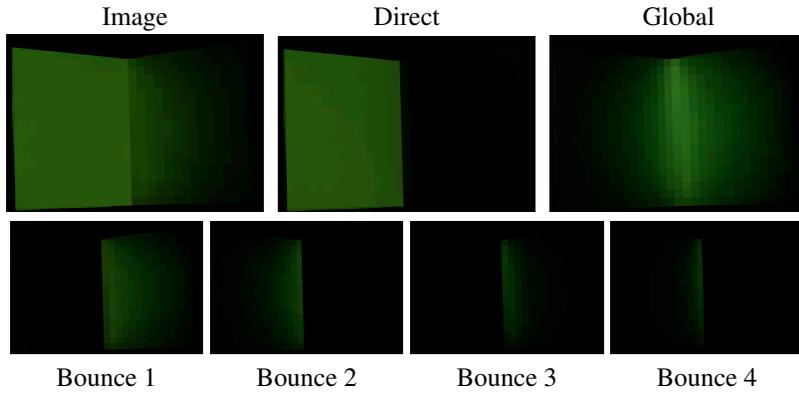


Figure 12: Separation of individual bounces. The scene is a white concave dihedral, with flat green projection on the left half. **Top row:** input image and separated direct and net global components. **Bottom row:** recovered indirect bounces. Note that successive bounces illuminate alternating walls of the dihedral, as expected.

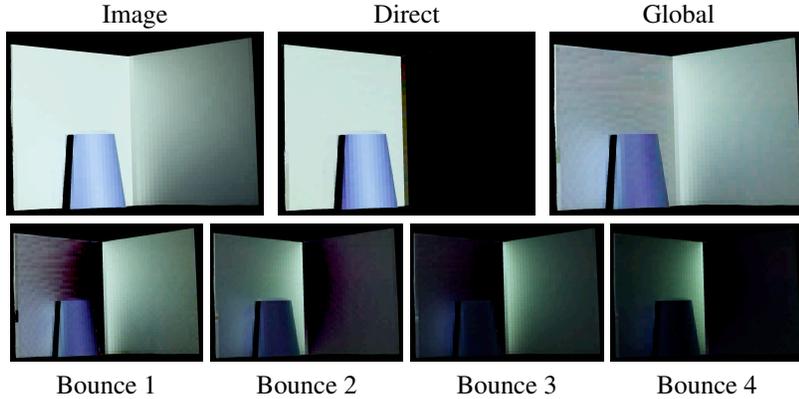


Figure 13: Bounce separation with occlusions and specularities. **Top row:** input image and separated direct and net global components. **Bottom row:** recovered indirect bounces. Note that successive bounces illuminate alternating walls and the specular highlight is present only in the direct component.

Thus, each successive run of our iterative inversion algorithm yields a bounce of light transport. Fig. 12 shows a didactic example demonstrating the accuracy of the bounce separation. The scene consists of a white dihedral with green light projected on the left half. Note that successive bounces of indirect illumination in the bottom row alternate perfectly between the two walls, as expected. Fig. 13 demonstrates the same with a non-Lambertian occluder present in the scene. We observe that the specular highlight is limited only to the direct component and absent from the indirect bounces, which is also expected.

This application is the same as [Seitz et al. 2005], but our algorithms are far more efficient. For instance, our iterative method recovers the direct component as well as each bounce of indirect illumination in 0.09 sec for the $4K \times 4K$ transport matrix in Fig. 13, while straightforward matrix inversion requires 4.6 sec. More importantly, our methods can efficiently operate on much higher resolution scenes that direct inversion cannot handle—for instance, Fig. 2 demonstrates bounce separation in a $131K \times 131K$ transport matrix. While an uncompressed matrix of that size cannot even be loaded in RAM, extrapolating from Fig. 10, a brute force inversion will require nearly 150 hours. In contrast, we require only 33ms per iteration in our (unoptimized) Matlab implementation, for a total of about 3 sec to separate each bounce. Note that the faster method of [Nayar et al. 2006] yields only the top row of Fig. 12 for a *particular* lighting

configuration, while we can separate all the bounces for *any* lighting, albeit at the expense of a more laborious acquisition.

10 Generality and Limitations

Finally, we briefly discuss the generality and limitations of the presented theory, especially in the context of our experimental setup in Section 9.

Choice of \mathbf{F} : It is important to note that \mathbf{F} need not correspond to the actual first bounce for an accurate light transport inversion. In numerical terms, the choice of $\mathbf{F} = \text{diag}(\mathbf{T})$ amounts to Jacobi preconditioning, which is convergent if \mathbf{T} is diagonally dominant. Thus, our choice of \mathbf{F} is valid for inversion of any light transport, even one arising from a non-Lambertian scene, as long as global effects do not dominate the transport. So, our theory is valid in its current form for applications like radiometric compensation in non-Lambertian scenes that rely merely on inversion of the light transport.

Non-Lambertian BRDFs: It is worth reemphasizing that our theoretical framework is derived in terms of the full light field and is applicable to light transports arising from complex BRDFs. In particular, our derivations remain true for non-Lambertian BRDFs, including anisotropic ones. Phe-

nomena such as translucency, subsurface and volumetric scattering, which cannot be modeled by an opaque BRDF, are not encompassed by this theory. Yet our results are well-behaved even in the presence of out-of-model effects like subsurface scattering, which are clearly visible in our experiments. An interesting avenue for future research is to explicitly incorporate such effects into our theory.

Single projector-camera setup: While our formulation is valid for any opaque BRDF when considering the full light field, our experiments employ a single projector and a single camera. Thus, an experimental setup like ours will necessitate the additional considerations of an operator \mathbf{P} that projects the output light field to the image and another operator \mathbf{Q} that raises the projector input to the full light field.

The acquired light transport can now be written as

$$\mathbf{T} = \mathbf{P}\mathbf{S}\mathbf{F}\mathbf{Q} \quad (52)$$

$$= \mathbf{P}(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots)\mathbf{F}\mathbf{Q} \quad (53)$$

$$= \mathbf{P}\mathbf{F}\mathbf{Q} + \mathbf{P}\mathbf{A}\mathbf{F}\mathbf{Q} + \mathbf{P}\mathbf{A}^2\mathbf{F}\mathbf{Q} + \dots \quad (54)$$

The direct lighting component in the observed image is $\mathbf{P}\mathbf{F}\mathbf{Q}$. In the case of the full light field, higher bounces are generated by a simple operator action \mathbf{A} . However, that is not true here due to the pre-multiplication by the projection operator \mathbf{P} , unless \mathbf{P} and \mathbf{A} commute multiplicatively (for which it is unlikely that any physical meaning exists). Thus, our physical interpretations in terms of bounces of light in Section 5 are only valid for the full light field, not for the single projector-camera setup of our experiments. Also, this makes the theory inexact for certain applications like bounce separation with a single projector-camera in non-Lambertian scenes.

But it is important to note that the theory does hold for a Lambertian scene even in the single projector-camera case. Since the camera direction is immaterial in that case, one need not consider \mathbf{P} for a radiometric analysis (or even \mathbf{Q} , if one ignores visibility and shadowing issues, as in [Seitz et al. 2005]). In practice, it is known that specular effects rapidly decay with bounces of interreflection (that is, higher bounces are increasingly diffuse), so the results obtained in our bounce separation experiments are still robust to moderate amounts of gloss.

Device limitations: We share some restrictions with other projector-camera systems, such as shutter speeds limited by projector refresh rates, color bleeding and non-linear color mixing ratios. For radiometric compensation, the projector cannot display negative values, which may lead to clipping artifacts in dark regions.

11 Conclusions and Future Work

The main contribution of this paper is a formulation of inverse light transport in computer vision, as a dual to the theory of forward rendering in computer graphics. This lends new insights for canceling interreflections in complex scenes, as well as fast computational methods for doing so. Our efficient algorithms, analogous to finite element radiosity and Monte Carlo path tracing in forward rendering, can handle transport resolutions far higher than previous methods.

From a theoretical perspective, we have just scratched the surface of analogies between forward and inverse methods. It is our hope that the framework of this paper forms the basis for discovering further insights into the structure of light transport and developing methods that couple fast acquisition and iterative inversion to perform radiometric compensation in dynamic scenes.

Acknowledgments: This work is funded by ONR YIP grant N00014-10-1-0032, ONR PECASE grant N00014-09-1-0741, a National Science Scholarship from A*STAR Graduate Academy of Singapore, as well as generous support from Adobe, NVIDIA, Intel and Pixar. We thank Joo Hwee Lim and Zhiyong Huang for kind support at I²R and anonymous reviewers of [Bai et al. 2010] for useful comments.

References

- ARVO, J., TORRANCE, K., AND SMITS, B. 1994. A framework for the analysis of error in global illumination algorithms. In *SIGGRAPH 94*, 75–84.
- BAI, J., CHANDRAKER, M., NG, T.-T., AND RAMAMOORTHY, R. 2010. A dual theory of inverse and forward light transport. In *European Conference on Computer Vision*.
- BIMBER, O., GRUNDHOEFER, A., ZEIDLER, T., DANCH, D., AND KAPAKOS, P. 2006. Compensating indirect scattering for immersive and semi-immersive projection displays. In *IEEE Virtual Reality*, 151–158.
- COHEN, M., AND WALLACE, J. 1993. *Radiosity and Realistic Image Synthesis*. Academic Press.
- DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *SIGGRAPH 00*, 145–156.
- DEMME, J. 1997. *Applied Numerical Linear Algebra*. SIAM.

- DING, Y., XIAO, J., TAN, K.-H., AND YU., J. 2009. Cata-dioptric projectors. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.
- FORSYTHE, G., AND LEIBLER, R. 1950. Matrix inversion by a Monte Carlo method. *Mathematical Tables and Other Aids to Computation* 4, 31 (jul), 127–129.
- FUJII, K., GROSSBERG, M., AND NAYAR, S. 2005. A Projector-camera System with Real-time Photometric Adaptation for Dynamic Environments. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.
- GARG, G., TALVALA, E., LEVOY, M., AND LENSCH, H. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *EuroGraphics Symposium on Rendering*, 251–262.
- GOLUB, G., AND VAN LOAN, C. 1996. *Matrix Computations*. John Hopkins University Press.
- GORTLER, S., SCHRÖDER, P., COHEN, M., AND HANRAHAN, P. 1993. Wavelet radiosity. In *SIGGRAPH 93*, 221–230.
- HANRAHAN, P., SALZMAN, D., AND AUPPERLE, L. 1991. A rapid hierarchical radiosity algorithm. In *SIGGRAPH 91*, 197–206.
- HASAN, M., PELLACINI, F., AND BALA, K. 2006. Direct to indirect transfer for cinematic relighting. *ACM Transactions on Graphics (SIGGRAPH 06)* 25, 3, 1089–1097.
- JENSEN, H. 2001. *Realistic Image Synthesis using Photon Mapping*. AK Peters.
- KAJIYA, J. 1986. The rendering equation. In *SIGGRAPH 86*, 143–150.
- MARSCHNER, S. 1998. *Inverse Rendering for Computer Graphics*. PhD thesis, Cornell University.
- MASSELUS, V., PEERS, P., DUTRE, P., AND WILLEMS, Y. 2003. Relighting with 4D incident light fields. *ACM Transactions on Graphics (SIGGRAPH 03)* 22, 3, 613–620.
- MUKAIGAWA, Y., KAKINUMA, T., AND OHTA, Y. 2006. Analytical compensation of inter-reflection for pattern projection. In *ACM VRST*, 265–268.
- NAYAR, S., PERI, H., GROSSBERG, M., AND BELHUMEUR, P. 2003. A Projection System with Radiometric Compensation for Screen Imperfections. In *Proceedings of IEEE International Workshop on Projector-Camera Systems*.
- NAYAR, S., KRISHNAN, G., GROSSBERG, M., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics (SIGGRAPH 06)* 25, 3, 935–944.
- NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2003. All-frequency shadows using non-linear wavelet lighting approximation. *ACM Transactions on Graphics (SIGGRAPH 03)* 22, 3, 376–381.
- NG, T.-T., PAHWA, R. S., BAI, J., QUEK, Q.-S., , AND TAN, K.-H. 2009. Radiometric Compensation Using Stratified Inverses. In *Proceedings of IEEE International Conference in Computer Vision*.
- PEERS, P., BERGE, K., MATUSIK, W., RAMAMOORTHI, R., LAWRENCE, J., RUSINKIEWICZ, S., AND DUTRE, P. 2006. A compact factored representation of heterogeneous subsurface scattering. *ACM Transactions on Graphics (SIGGRAPH 06)* 25, 3, 746–753.
- RAMAMOORTHI, R., AND HANRAHAN, P. 2001. A signal-processing framework for inverse rendering. In *SIGGRAPH 01*, 117–128.
- RASKAR, R., WELCH, G., LOW, K., AND BANDYOPADHYAY, D. 2001. Shader lamps. In *EuroGraphics Workshop on Rendering*.
- SATO, I., SATO, Y., AND IKEUCHI, K. 1999. Illumination distribution from shadows. In *CVPR*, 1306–1312.
- SCHRÖDER, P., AND HANRAHAN, P. 1993. On the form factor between two polygons. In *SIGGRAPH 93*, 163–164.
- SEITZ, S., MATSUSHITA, Y., AND KUTULAKOS, K. 2005. A theory of inverse light transport. In *Proceedings of IEEE international conference in Computer Vision*, 1440–1447.
- SEN, P., CHEN, B., GARG, G., MARSCHNER, S., HOROWITZ, M., LEVOY, M., AND LENSCH, H. 2005. Dual Photography. *ACM Transactions on Graphics (SIGGRAPH 05)* 24, 3, 745–755.
- SLOAN, P., KAUTZ, J., AND SNYDER, J. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM Transactions on Graphics (SIGGRAPH 02)* 21, 3, 527–536.
- VEACH, E. 1998. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University.
- WETZSTEIN, G., AND BIMBER, O. 2007. Radiometric Compensation through Inverse Light Transport. In *Proceedings of Pacific conference on computer graphics and applications*, 391–399.
- YU, Y., DEBEVEC, P., MALIK, J., AND HAWKINS, T. 1999. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *SIGGRAPH 99*, 215–224.