

硕 士 研 究 生 读 书 报 告



题目 Regression-based Monte Carlo Integration

作者姓名 朱泽华

作者学号 22251034

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二○二二年十二月二十五日

Regression-based Monte Carlo Integration[1]

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Li Qilei

By

Zhu Zehua

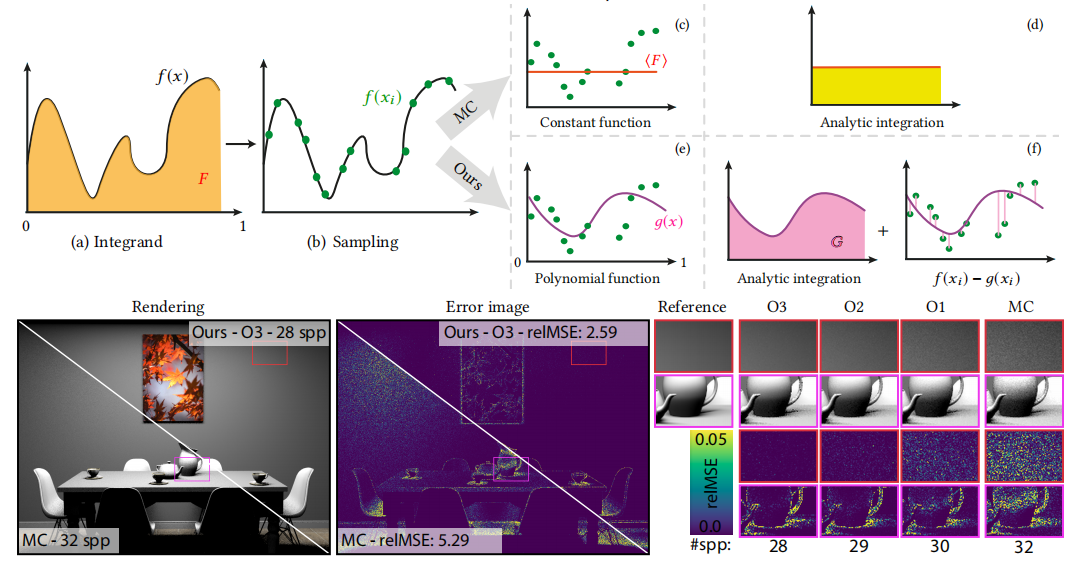
Zhejiang University, P.R. China

2022.12.25

1引言

在现代的光追任务中，我们想了很多办法来优化光追的性能，并且基于光追需要通过蒙特卡洛方法计算积分，提出了两种主要的优化方法，分别是控制变量和重要性采样。后来的学者基于这两种方法，又提出了许多进一步的优化方案。而本文正是发表在siggraph2022上的一篇文章，关于如何利用回归分析来达到优化MC积分的效果

在这篇文章中，主要被用到的方法就是回归、多项式和控制变量。在原本广泛应用的蒙特卡洛积分的基础上倒退一步，通过回归拟合出一个和被积函数相似的可以求出解析解的函数，在这个基础上用 Control variates (CVs) 的思路，在采样数相同的情况下可以显著降低均方差，提高积分器的效率。



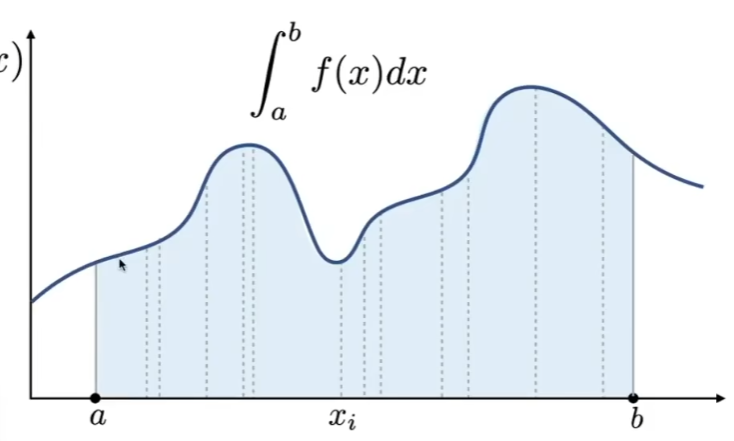
上图是这个优化方案的最终效果，可以看到他的基本思路以及最终的效果还是比较理想的。基于回归的MC积分，把多项式作为模型函数，Ox中的x代表x阶多项式，通过回归找到一个对应阶数的多项式。

以往的一些研究一般会专注于通过寻找一个模型函数，通过模型函数的拟合寻找控制变量，最终达到优化的效果。但是不管是这篇文章，还是之前的研究者进行的各项工作都不能最终实现对各种场景下的高效优化。

下面我们先来详细了解一下这篇文章所做的贡献。因为他开创了通过多项式来直接拟合出控制变量的新方法，这个新思路可能可以引出很多的新工作，所以可以作为一篇引路的文章带大家看一看现在最新的一些光追研究人员所作的工作以及基于这篇文章的提出，未来我们可以进行哪些方向的研究，并在此基础上提出一些更加通用且高效的方案。

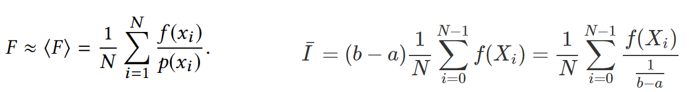
我觉得，现在的离线渲染和光追优化这个部分其实已经达到了比较不错的效果，目前的各种优化和进一步改进都是为了能够在渲染中达到更加真实的效果。如果有一个新的大方向能被提出，并且被验证有着不错的效果，是会被学界大力推崇的。

**2 Monte Carlo method**



蒙特卡罗方法（Monte Carlo method），也称统计模拟方法，是1940年代中期由于科学技术的发展和电子计算机的发明，而提出的一种以概率统计理论为指导的数值计算方法。是指使用随机数（或更常见的伪随机数）来解决很多计算问题的方法。

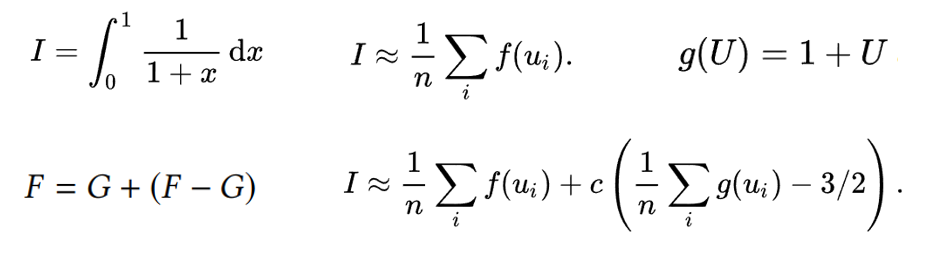
在图形学的ray tracing中，主要的用途就是通过采样来估计一个无法求出解析解的定积分的值。主要公式如下：



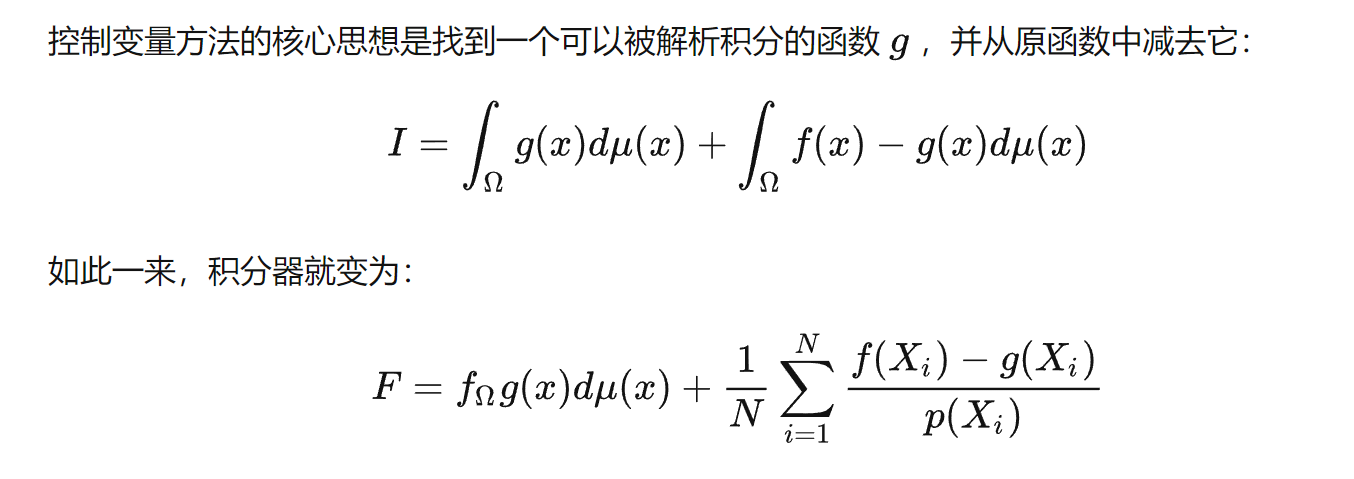
通过这个式子可以看出，我们的估算值是一个无偏估计，只要采样次数达到一定数量，就能求出一个近似解。

**3 Control variates控制变量方法**

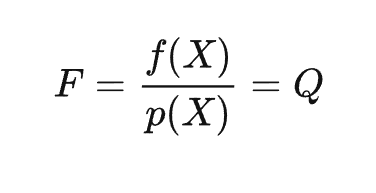
控制变量方法是一个应用在MC中的优化方法，用来降低MC积分过程中因为采样而产生的方差。它利用有关已知量估计误差的信息来减少未知量估计误差。下面的图中展示了一个相关的简单例子：



通过合理的选取 g ，这个新的积分器将比原来的拥有更低的方差。该方法可以与重要性采样联合使用，但是需要注意的是，如果 g∝p ，则方差并不会变低。CVs是一种可以用来降低MC积分计算方差的方法，如果选择的控制变量g和原被积函数f越近似，这个方法的效果就越好。g必须是一个存在积分解析解，可以直接计算出定积分值的函数

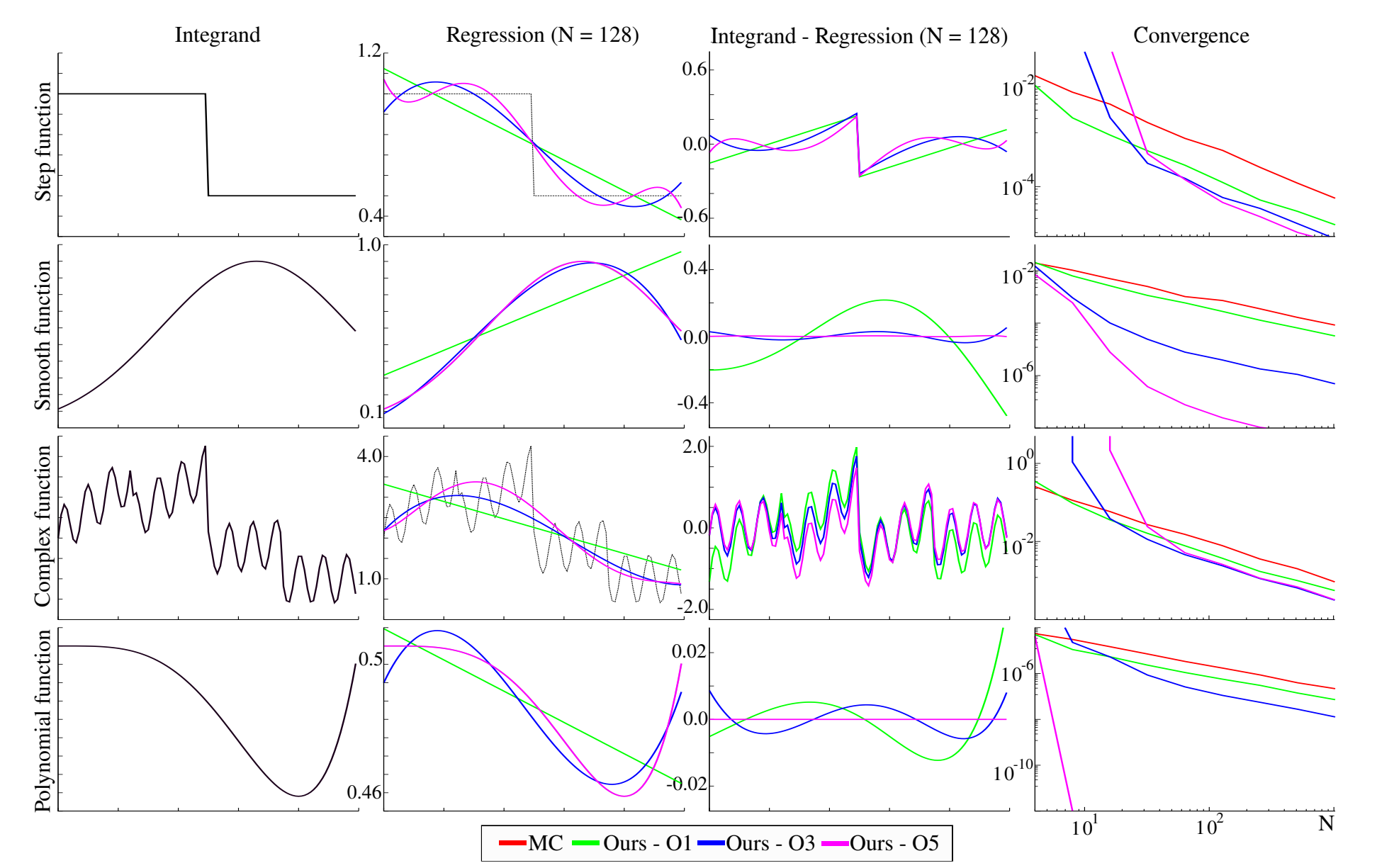


和他类似的另一个概念，重要性采样的方法也是类似选取一个函数的思路。重要性采样(Importance sampling)：通过选取一个与被积函数近似的密度函数 p ，可以有效降低方差。最佳方案就是让 p(x)=1/Qf(x) ，则该积分器就会完全摆脱方差。因为对于所有的采样点：



CVs这个东西，随便选一个函数g不一定能减小最终的方差，能减小多少还取决于f和g的相似程度，这篇文章的主要贡献是，证明了最小二乘回归不仅能帮我们自动找到这个g，并且如果回归得到的结果里包含了一个常数时，他还能减小期望误差。

如下图我们可以看到，即使是回归得到的函数和原被积函数没有那么相似，比如第一个和第三个，在采样次数比较多之后，误差都能小于一般的MC积分。如果得到的函数拟合的比较好的话，从图上可以看到，方差会快速降低。



上图第一列是被积函数，第二列不同颜色的线分别代表用不同阶数的多项式回归得到的结果我就把他称为g好了，第三列是拿原函数f减去g得到的图像。最后一列是采样次数和方差与普通mc积分的比较

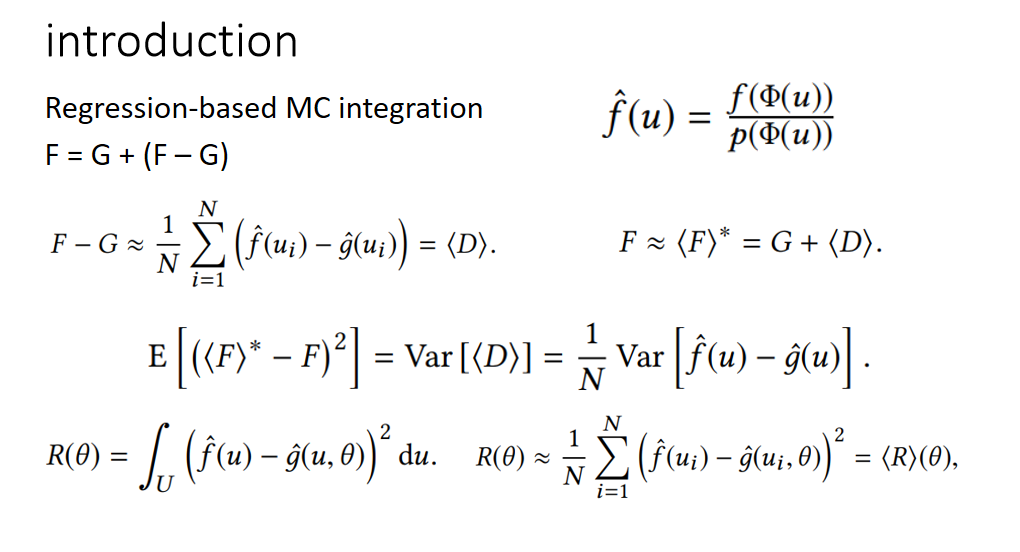
所以如果很不幸的是回归返回了一个常数c，可以证明用了CVs之后，最坏情况就是跟一般MC积分的结果一样，这样就相当于浪费了计算回归的时间，但是结果不会比普通MC积分要差。

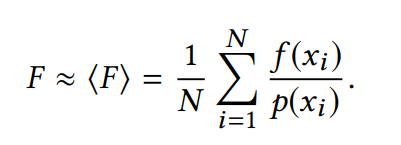
**4回归分析**

这篇文章并没有提出一种分析回归的新方法，他们的主要方案是最简单的计算残差，通过找残差最小的函数找到那个和当前阶数的多项式最为相似的函数作为控制变量的解。按照他的方法，再结合前面提到的引入一个常数c，它只需要通过回归确定残差最小的CVs，就可以保证至少不会比普通的蒙特卡洛积分要差。

对于用最小二乘法的回归分析，在实践中，由于最小二乘回归的解对于任何有限 N 都不是精确的，因此对于任何有限 N 都不能保证这种减少误差的理论特性。尽管如此，我们的数值结果表明，即使在 N 的实际范围内，误差也会减少并且稳定。

之前的研究者在用控制变量进行蒙特卡洛积分的时候，1.不会引入常数，因为他们觉得这样并不会让方差减小。2.如果发现引入控制变量让结果方差更差了，就选择用一般的蒙特卡洛积分解决。

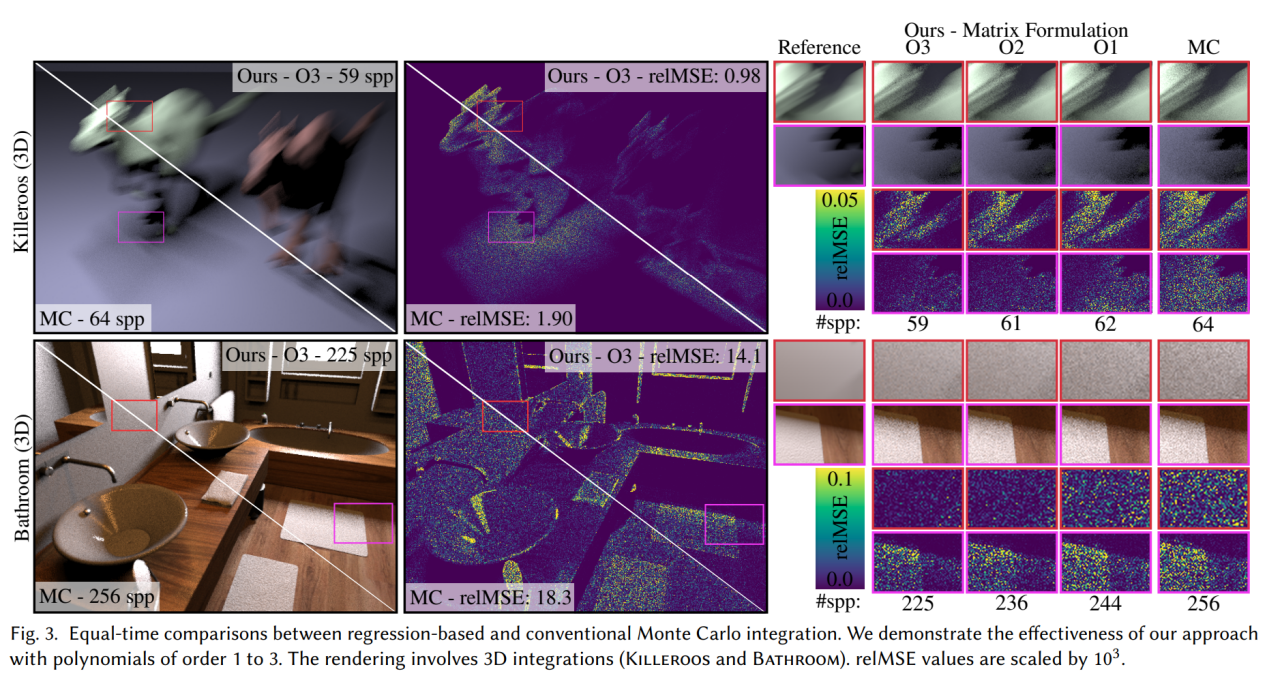


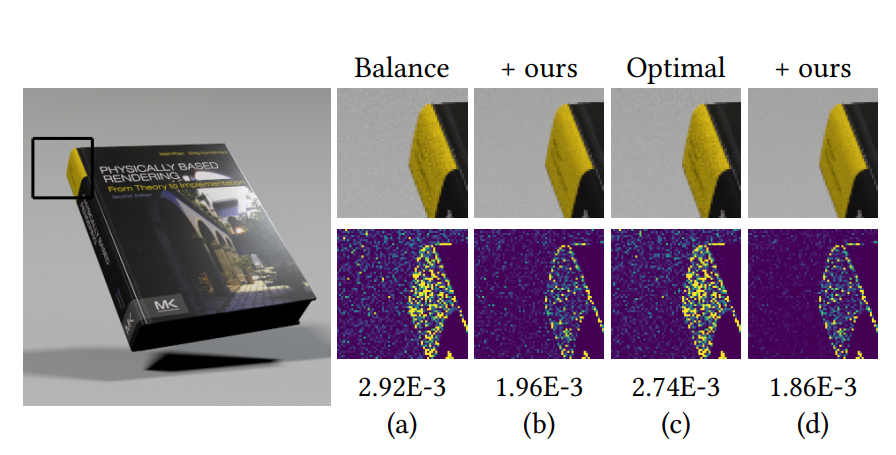


如上图所示，这样找到的控制变量，只需要找到和定义的多项式残差最小的的函数g就是当前能让这个积分的方差达到最小的情况。这里F代表初始函数f的积分，引入控制变量后，我们可以根据MC得到F的近似

R(θ)是两个函数f和g的残差，但是因为考虑到残差这个积分大多数情况下是不可积的，所以再应用一次mc，就能推出，F和F\*之间的方差近似于f和g的残差。

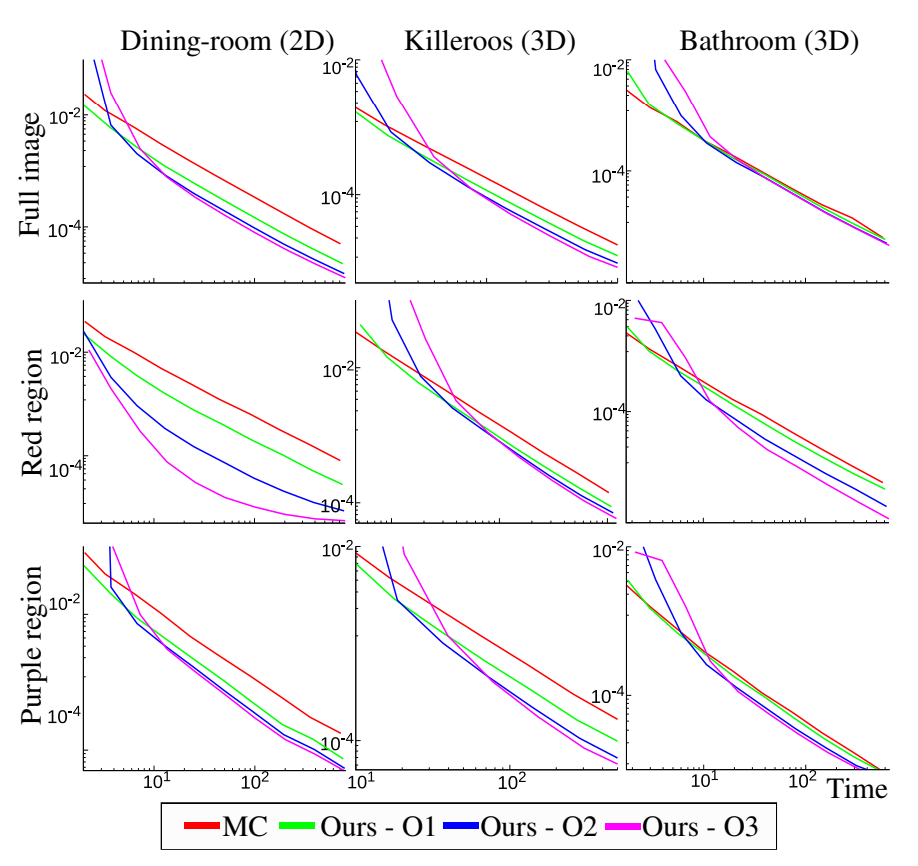
**5 小结**



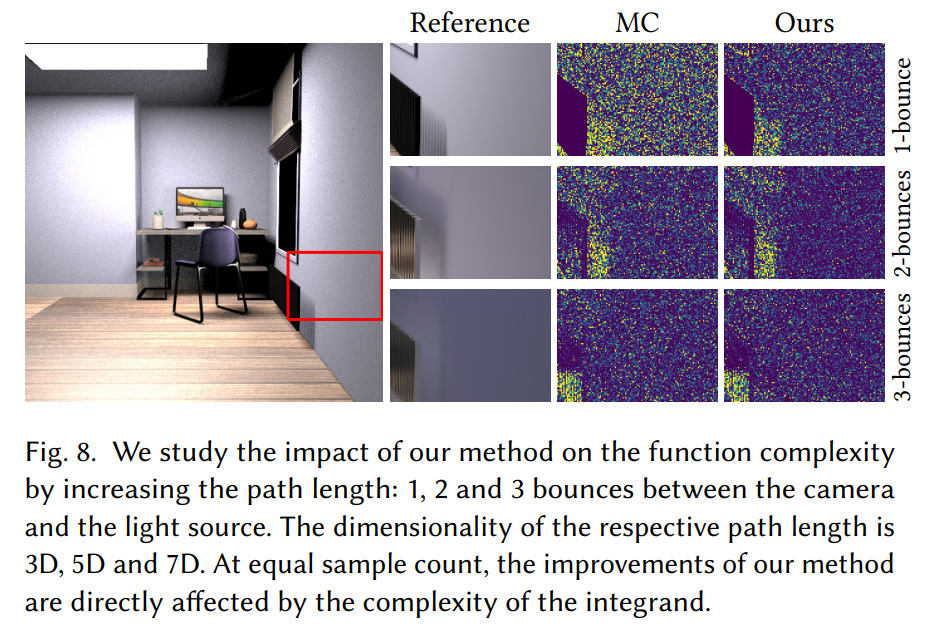
和前面做的比较好的工作相比，这篇文章减少的误差在整个场景中比较均匀。

在采样次数足够多的时候，这种方法的蒙特卡洛积分不会比一般mc积分效果要差。

As future work, Crespo et al.’s adaptive model for дˆ can potentially be useful in our formulation, but developing a practical least-squares regression algorithm for piecewise polynomials remains challenging. 未来工作可以结合Crespo等人的工作，实现自适应的分段多项式

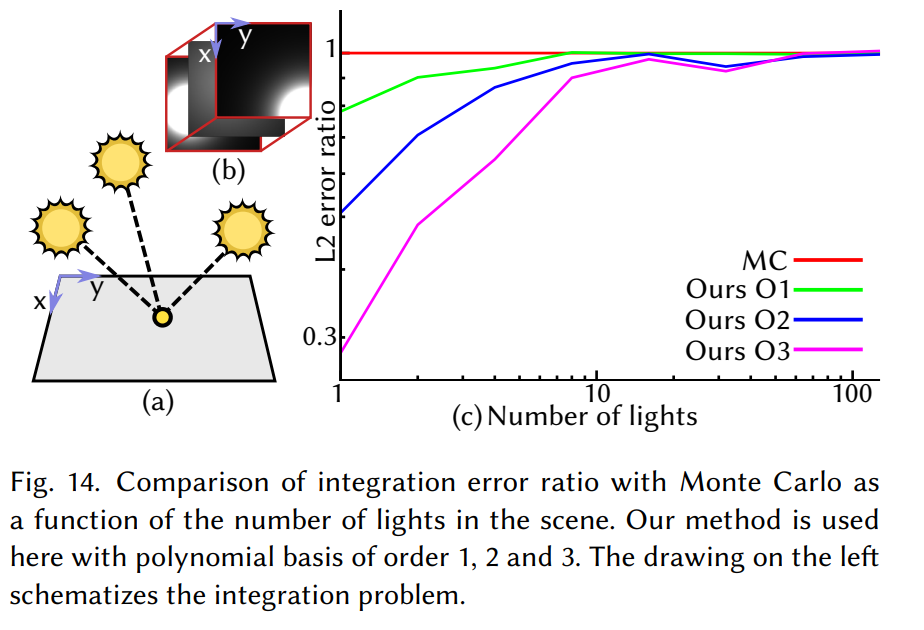


从测试结果看，好像也没有提升特别多，有些场景在采样次数不够的情况下误差还有所增加。



如上图可以看到，这种优化方法对光源直射的场景优化效果最好，如果是光线传播了多次的，效果就不那么明显了。

As long as a constant function is included in least-squares regression, our method would be no worse than MC integration, no matter, what basis functions are chosen. 只要最小二乘回归里包含常数，无论选择什么基函数，都不会比MC积分差。



多光源场景也是这样，随着光源增加，优化效果就逐渐趋近于一般mc积分了，因为多项式不能很好的表示不连续的函数。

这里是给出了多项式和普通mc积分的比较，如果灯光数量不多，用多项式能明显优于一般mc，因为多项式基要表示不连续函数很困难，所以光源增加后，多项式方法也就趋于一般mc方法了。

固定了光线传播的次数，因为回归需要确定搜索空间维度。

所以综上所述，这篇文章的主要贡献是，提出了可以把多项式、回归分析、控制变量结合起来。

参考文献

1. Salaün C, Gruson A, Hua B S, et al. Regression-based Monte Carlo integration[J]. ACM Transactions on Graphics (TOG), 2022, 41(4): 1-14.