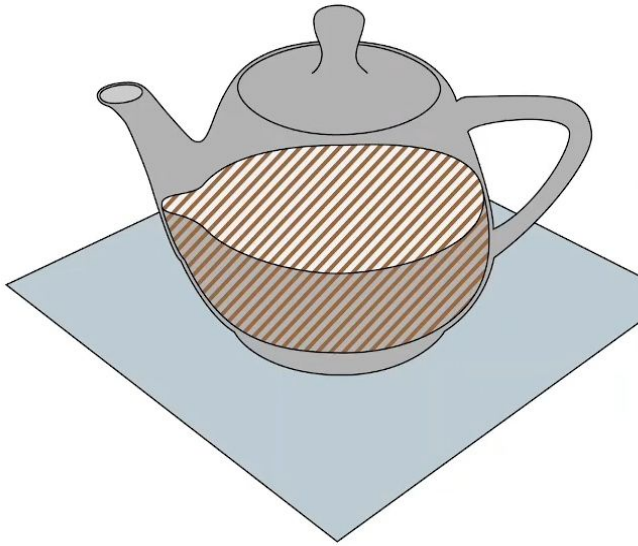


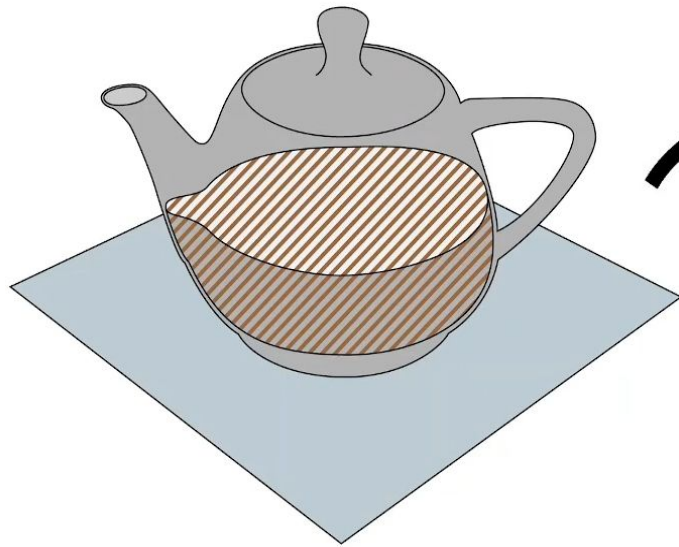
Recursive Control Variates for Inverse Rendering

Baptiste Nicolet Fabrice Rousselle Jan Novák Alexander Keller Wenzel Jakob Thomas Müller
SIGGRAPH 2023

Inverse rendering



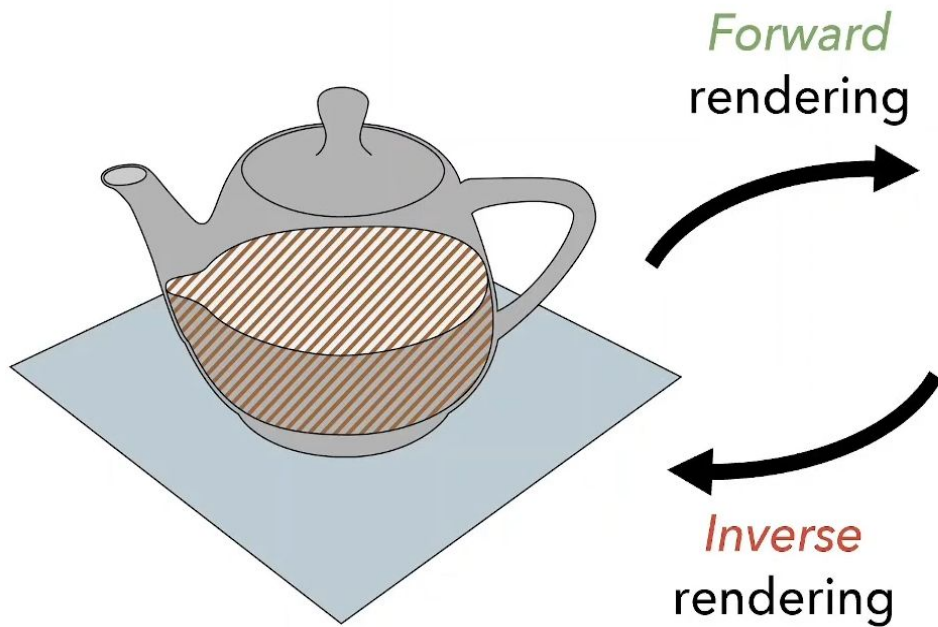
Inverse rendering



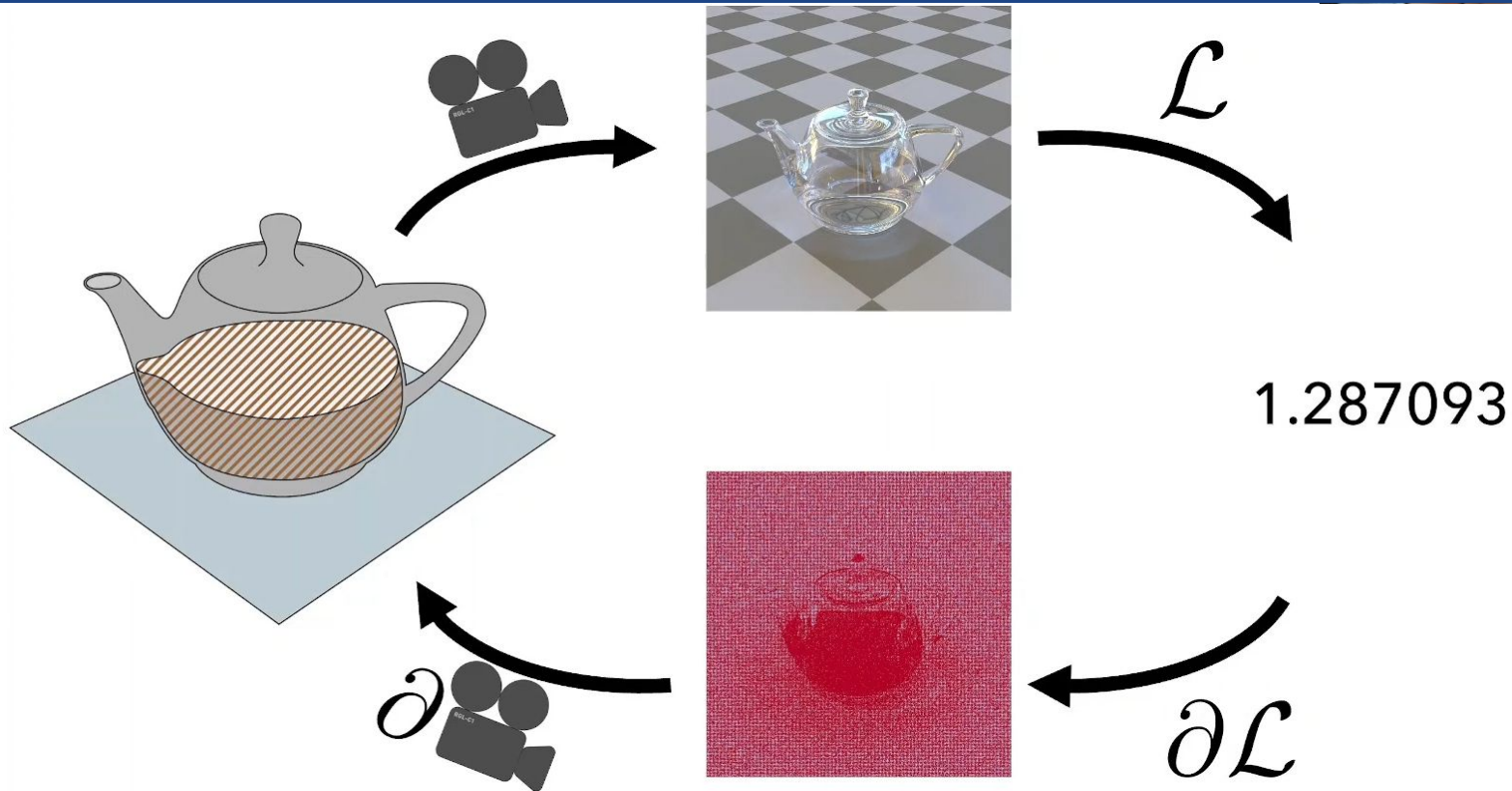
Forward
rendering



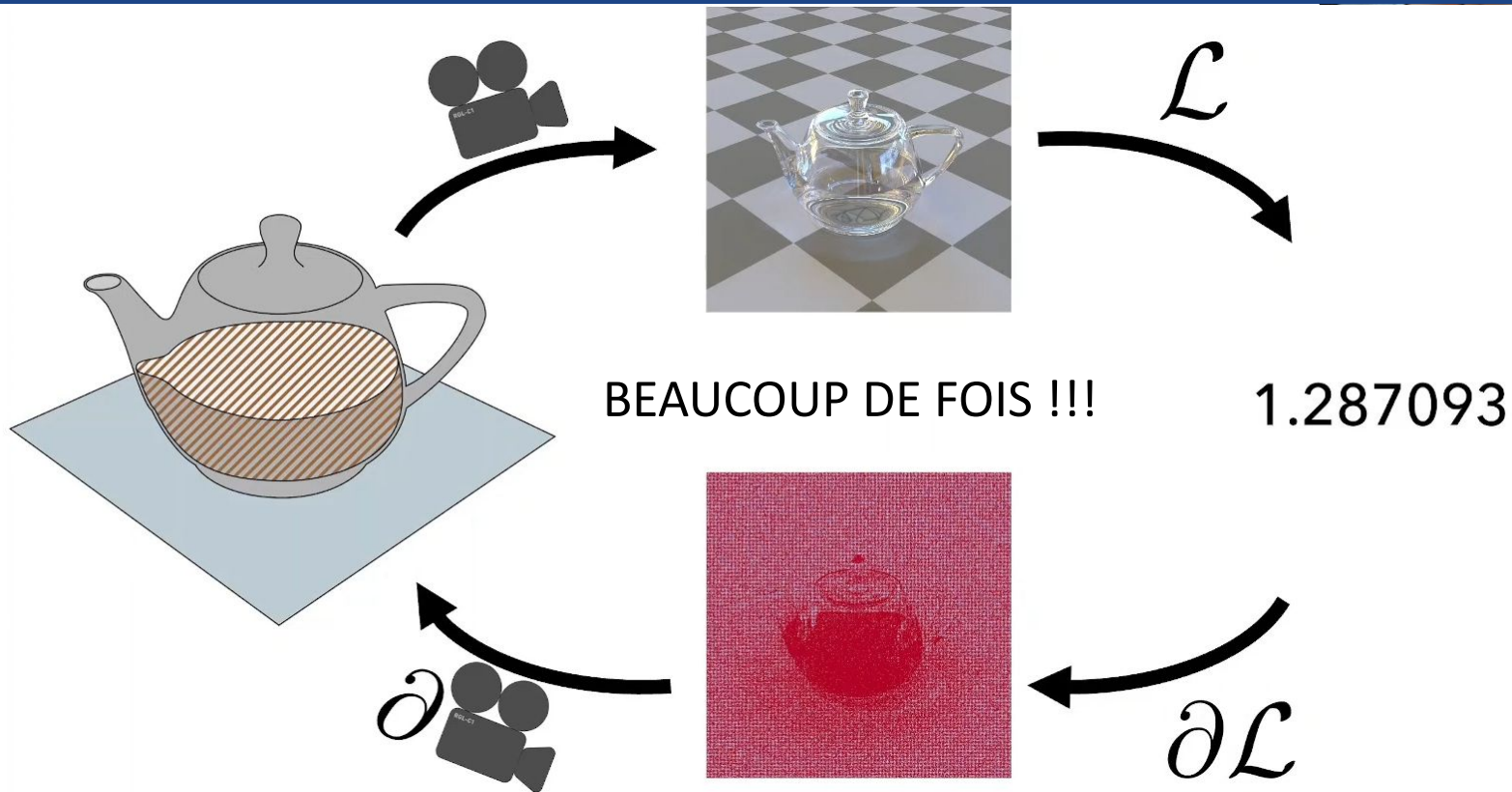
Inverse rendering



Inverse rendering



Inverse rendering

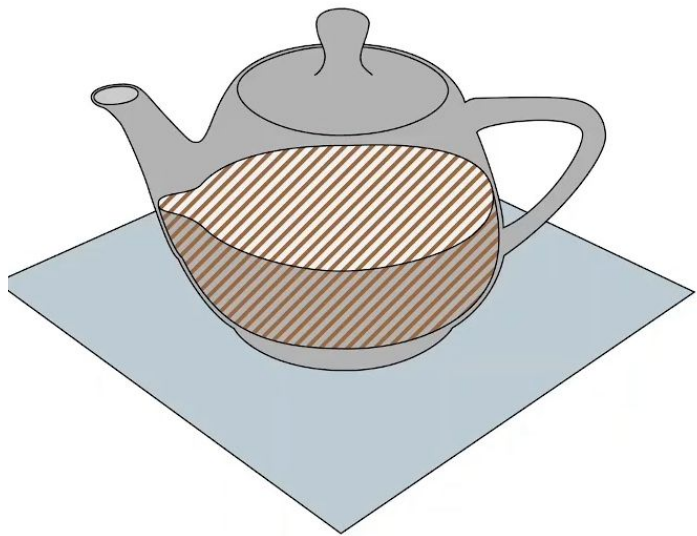


Inverse rendering



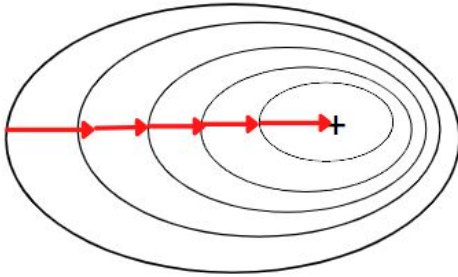
Inverse rendering

Rendering is noisy !

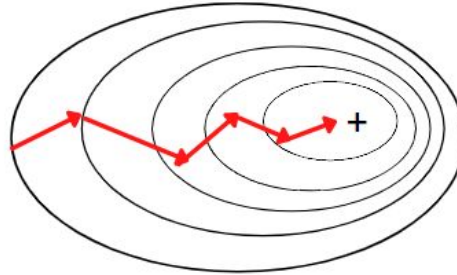


Noisy gradients is ok !

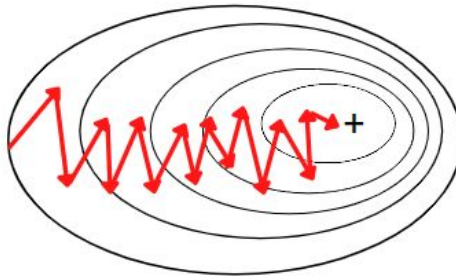
Batch Gradient Descent



Mini-Batch Gradient Descent



Stochastic Gradient Descent



Note: gradients are biased

On cherche (par descente de gradient):

$$\theta^* = \arg \min \mathcal{L}(F_\theta, I)$$

On estime la dérivée par chain-rule

$$\begin{aligned}\mathbb{E} [\partial_\theta (F_\theta, I)] &= \mathbb{E} [\partial_\theta F_\theta \cdot \partial_F \mathcal{L}(F_\theta, I)] \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \mathbb{E} [\partial_F \mathcal{L}(F_\theta, I)] \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \partial_F \mathcal{L}(\mathbb{E} [F_\theta], I)\end{aligned}$$

Note: gradients are biased

On cherche (par descente de gradient):

$$\theta^* = \arg \min \mathcal{L}(F_\theta, I)$$

On estime la dérivée par chain-rule

Ok seulement si indépendance

$$\begin{aligned}\mathbb{E} [\partial_\theta (F_\theta, I)] &= \mathbb{E} [\partial_\theta F_\theta \cdot \partial_F \mathcal{L}(F_\theta, I)] \quad \text{red arrow} \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \mathbb{E} [\partial_F \mathcal{L}(F_\theta, I)] \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \partial_F \mathcal{L}(\mathbb{E} [F_\theta], I)\end{aligned}$$

Note: gradients are biased

On cherche (par descente de gradient):

$$\theta^* = \arg \min \mathcal{L}(F_\theta, I)$$

On estime la dérivée par chain-rule

Ok seulement si indépendance

$$\begin{aligned} \mathbb{E} [\partial_\theta (F_\theta, I)] &= \mathbb{E} [\partial_\theta F_\theta \cdot \partial_F \mathcal{L}(F_\theta, I)] \\ &\stackrel{\text{Ok seulement si dL est linéaire (MSE)}}{=} \mathbb{E} [\partial_\theta F_\theta] \cdot \mathbb{E} [\partial_F \mathcal{L}(F_\theta, I)] \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \partial_F \mathcal{L}(\mathbb{E} [F_\theta], I) \end{aligned}$$

Note: gradients are biased

On cherche (par descente de gradient):

$$\theta^* = \arg \min \mathcal{L}(F_\theta, I)$$

On estime la dérivée par chain-rule

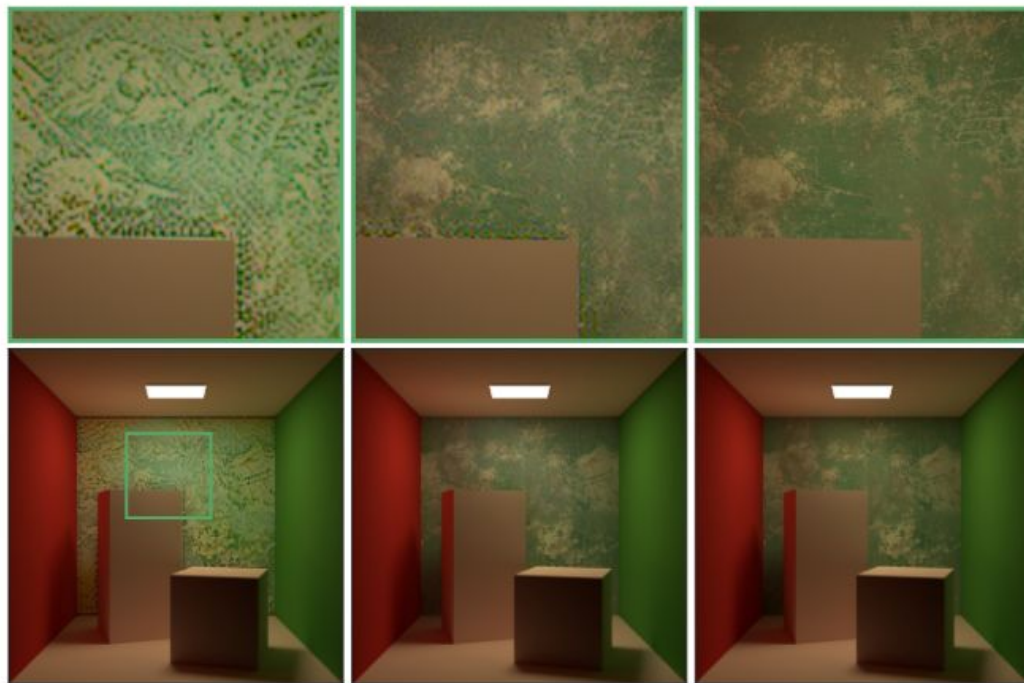
[Vicini 2021]

Ok seulement si indépendance

$$\begin{aligned} \mathbb{E} [\partial_\theta (F_\theta, I)] &= \mathbb{E} [\partial_\theta F_\theta \cdot \partial_F \mathcal{L}(F_\theta, I)] \\ &\stackrel{\text{Ok seulement si dL est linéaire (MSE)}}{=} \mathbb{E} [\partial_\theta F_\theta] \cdot \mathbb{E} [\partial_F \mathcal{L}(F_\theta, I)] \\ &= \mathbb{E} [\partial_\theta F_\theta] \cdot \partial_F \mathcal{L}(\mathbb{E} [F_\theta], I) \end{aligned}$$

Dans le papier : on estime
le biais et on l'enlève
(ok si peu de variance)

Note: gradients are biased

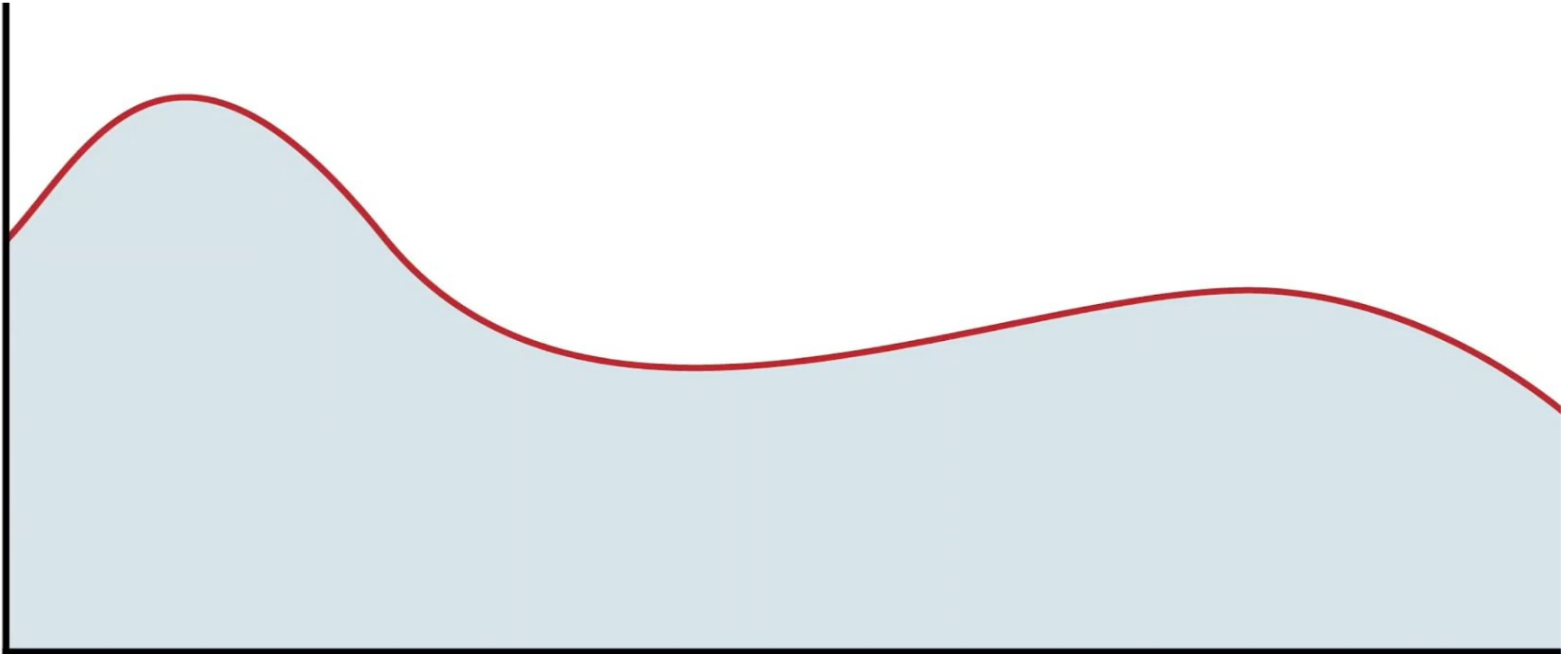


(a) Baseline

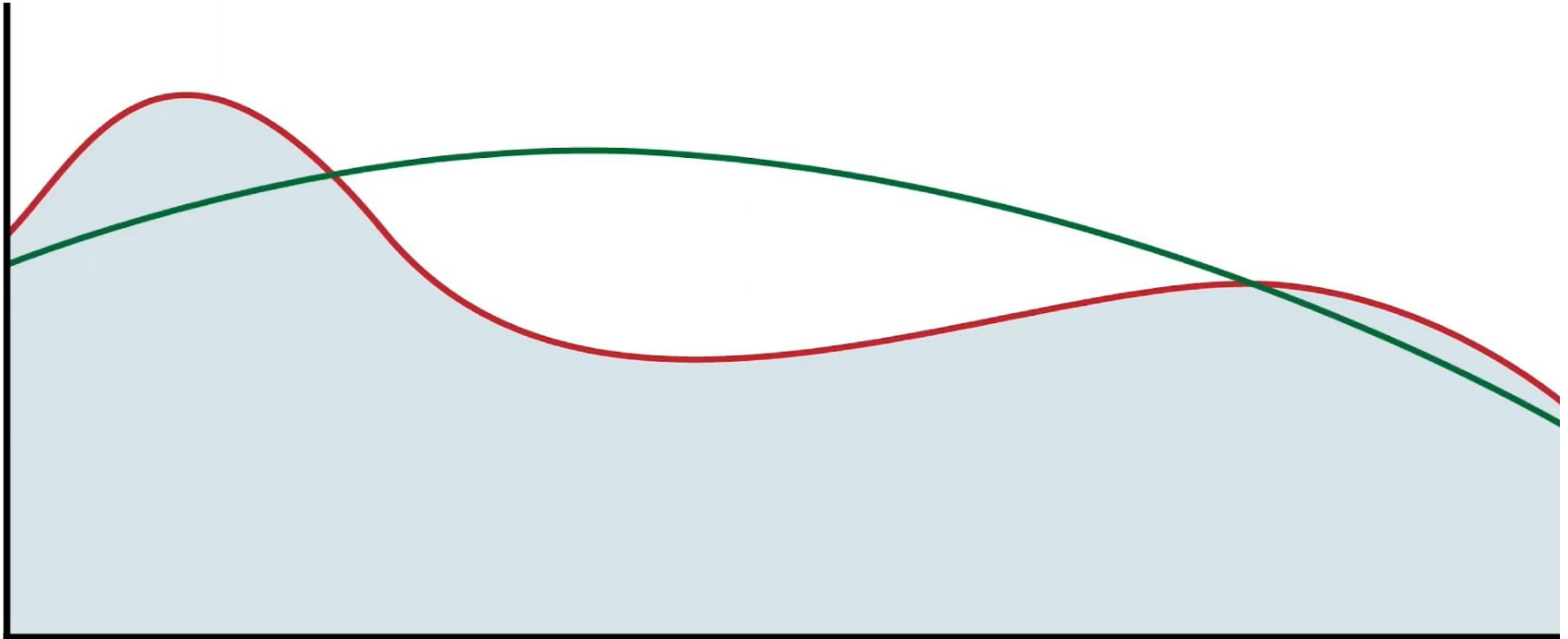
(b) Ours

(c) Reference

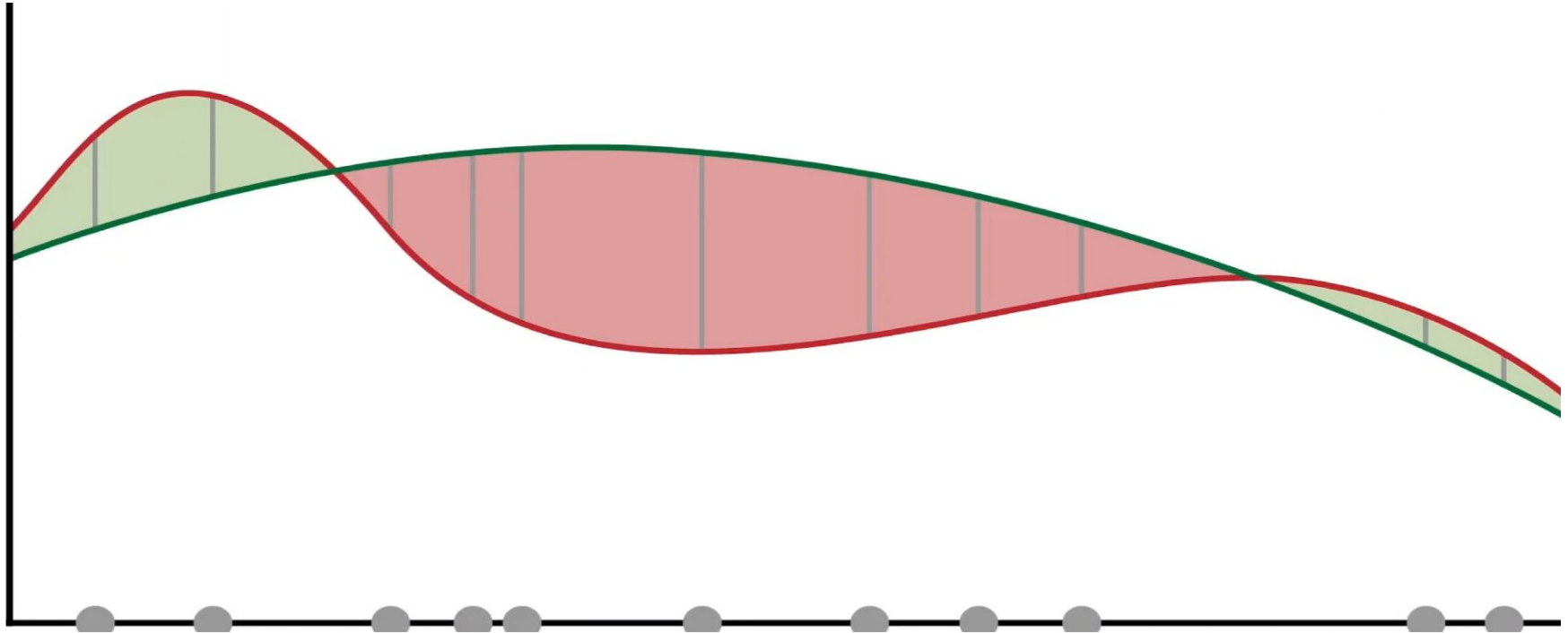
Control variates



Control variates



Control variates



Control variates (maths)

Mathématiquement :

$$\langle F_{CV} \rangle = G + \langle F \rangle - \langle G \rangle = \langle F \rangle + (G - \langle G \rangle)$$

Souvent, on ne connaît qu'une forme ressemblante :

$$\langle F_{CV} \rangle = \langle F \rangle + \alpha(G - \langle G \rangle)$$

Note: tous les choix de alpha sont valides (n'introduisent pas de biais). On a quand même :

$$\alpha_{opt} = \frac{Cov [\langle F \rangle , \langle G \rangle]}{Var [\langle G \rangle]}$$

Control variates (maths)

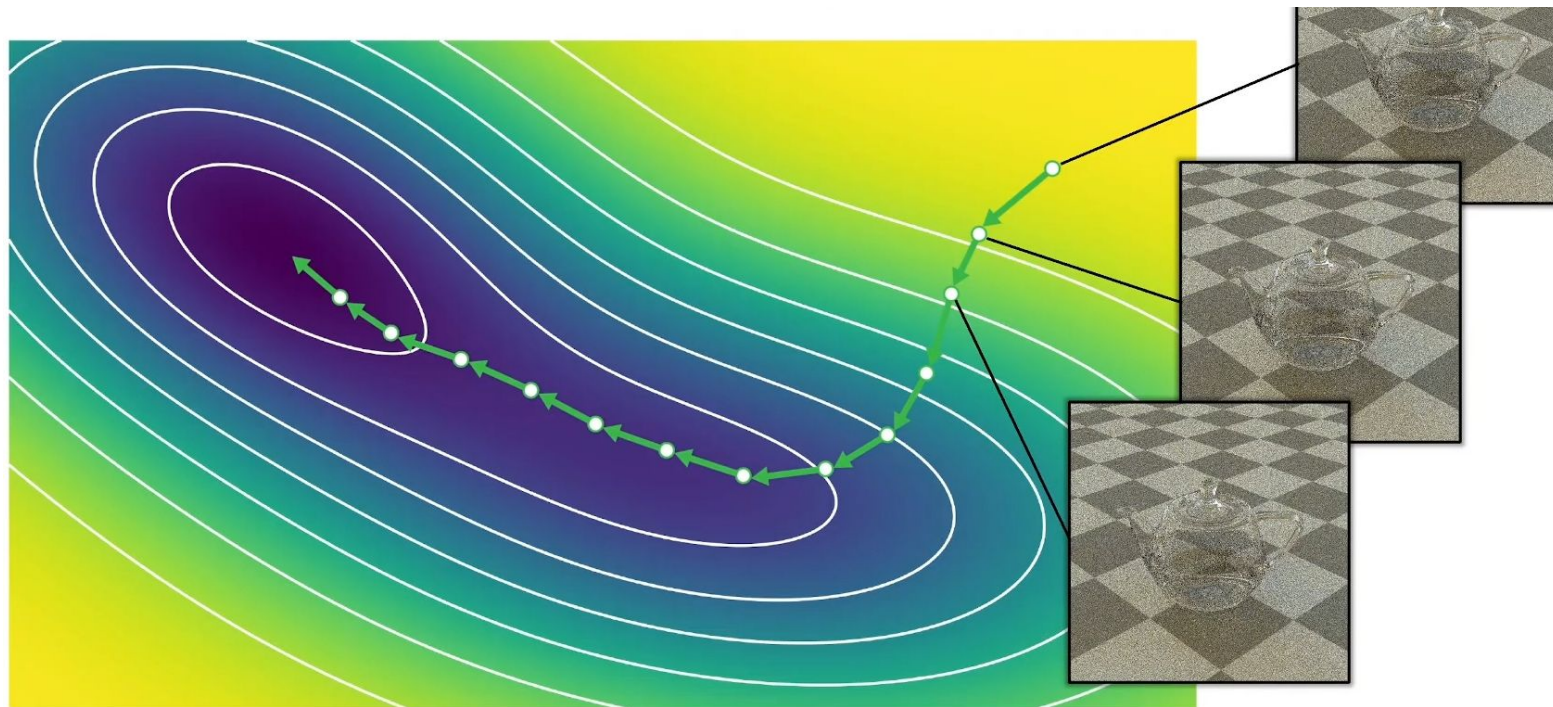
Si on ne connaît pas G :

$$\langle F_{CV} \rangle = \langle F \rangle + \alpha(\langle G \rangle - \langle G \rangle_0)$$

Et le alpha optimal devient:

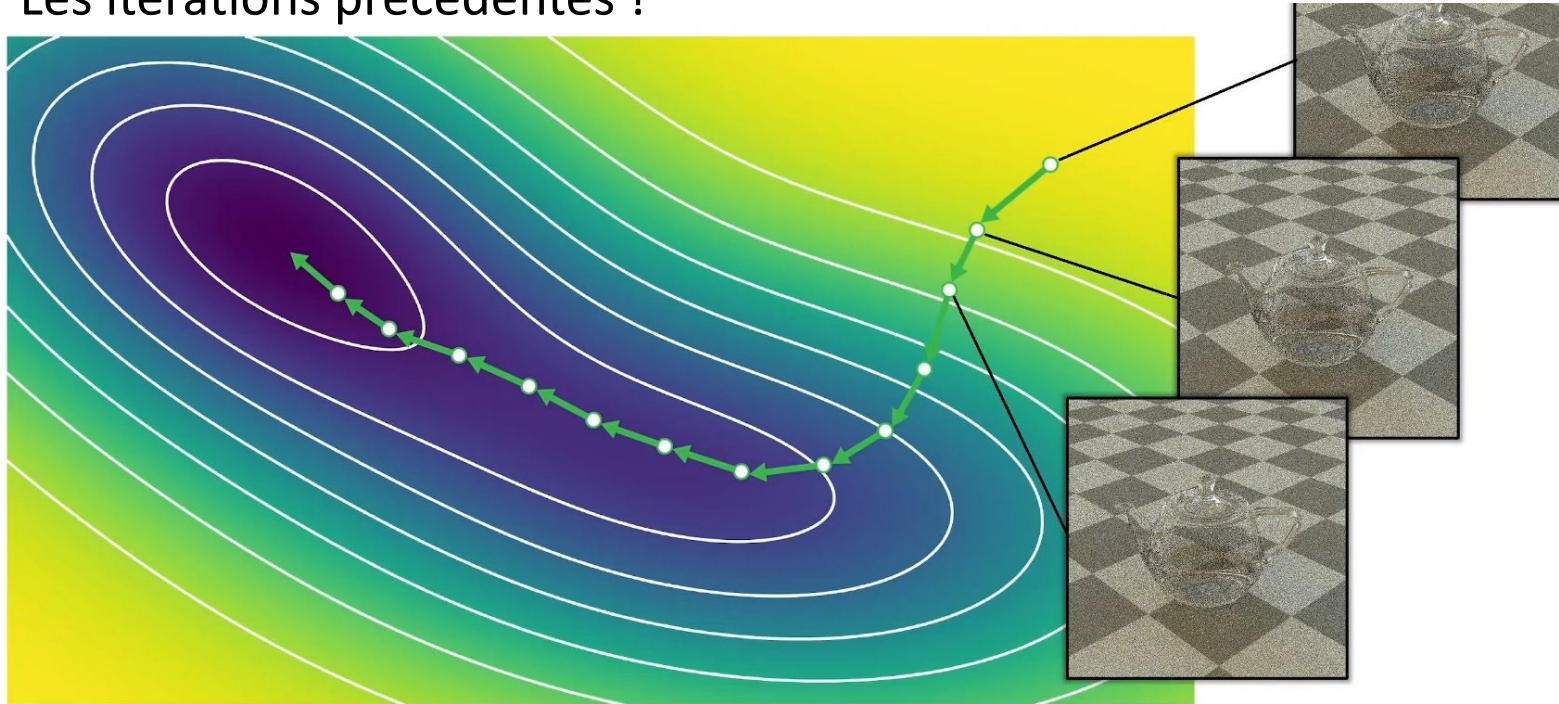
$$\alpha_{opt} = \frac{Cov [\langle F \rangle , \langle G \rangle - \langle G \rangle_0]}{Var [\langle G \rangle - \langle G \rangle_0]}$$

Que prendre pour G ?



Que prendre pour G ?

Les itérations précédentes !



Recursive control variates :

Rousselle et al. [2016] propose to use the original rendering $\langle F_0 \rangle_p$ as a control variate, resulting in the estimator

$$\langle F_1 \rangle_{CV} = \langle F_1 \rangle_1 + \alpha (\langle F_0 \rangle_0 - \langle F_0 \rangle_1), \quad (4)$$

We generalize the concept in Equation (4) to a sequence, where the new estimator

$$\langle F_n \rangle_{CV} = \langle F_n \rangle_n + \alpha_n (\langle F_{n-1} \rangle_{CV} - \langle F_{n-1} \rangle_n) \quad (6)$$

Quel poids prendre ?

The starting point of this sequence is set to

$$\langle F_0 \rangle_{CV} := \langle F_0 \rangle_0 \quad (7)$$

and its optimal control weight is given by

$$\alpha_n = \frac{\text{Cov}[\langle F_n \rangle_n, \langle F_{n-1} \rangle_n]}{\text{Var}[\langle F_{n-1} \rangle_n] + \text{Var}[\langle F_{n-1} \rangle_{CV}]} . \quad (8)$$

Quel poids prendre ?

The starting point of this sequence is set to

$$\langle F_0 \rangle_{CV} := \langle F_0 \rangle_0 \quad (7)$$

and its optimal control weight is given by

$$\alpha_n = \frac{\text{Cov}[\langle F_n \rangle_n, \langle F_{n-1} \rangle_n]}{\text{Var}[\langle F_{n-1} \rangle_n] + \text{Var}[\langle F_{n-1} \rangle_{CV}]} . \quad (8)$$

Pas un optimal sur toute la séquence en réalité ! : il faudrait le calculer sur toutes les itérations précédentes ce qui nécessite un coût mémoire trop élevé...

Quel poids prendre ?

$$\alpha_n = \frac{\text{Cov}[\langle F_n \rangle_n, \langle F_{n-1} \rangle_n]}{\text{Var}[\langle F_{n-1} \rangle_n] + \text{Var}[\langle F_{n-1} \rangle_{CV}]}$$

Alpha joue le rôle de contrôle entre le passé et le courant :

- Si il y a eu un “breaking change”: alpha = 0 et seules les itérations actuelles compte
- Si il n’y a pas eu de changement : alpha = n/n+1 et on obtient une moyenne pondérée

Comment estimer les paramètres ?

Welford's algorithm:
$$\left(\frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 \right) \cdot \frac{n}{n-1}$$

Attention: la définition de FCV violent les hypothèses d'indépendance pour calculer la variance. On utilise alors :

$$\begin{aligned} \text{Var}[\langle F_n \rangle_{\text{CV}}] &= \text{Var}[\langle F_n \rangle_n + \alpha_n (\langle F_{n-1} \rangle_{\text{CV}} - \langle F_{n-1} \rangle_n)] & (10) \\ &= \text{Var}[\langle F_n \rangle_n] + \alpha_n^2 \text{Var}[\langle F_{n-1} \rangle_n] \\ &\quad - 2\alpha_n \text{Cov}[\langle F_n \rangle_n, \langle F_{n-1} \rangle_n] + \alpha_n^2 \text{Var}[\langle F_{n-1} \rangle_{\text{CV}}] \end{aligned}$$

Results



It 0

Results



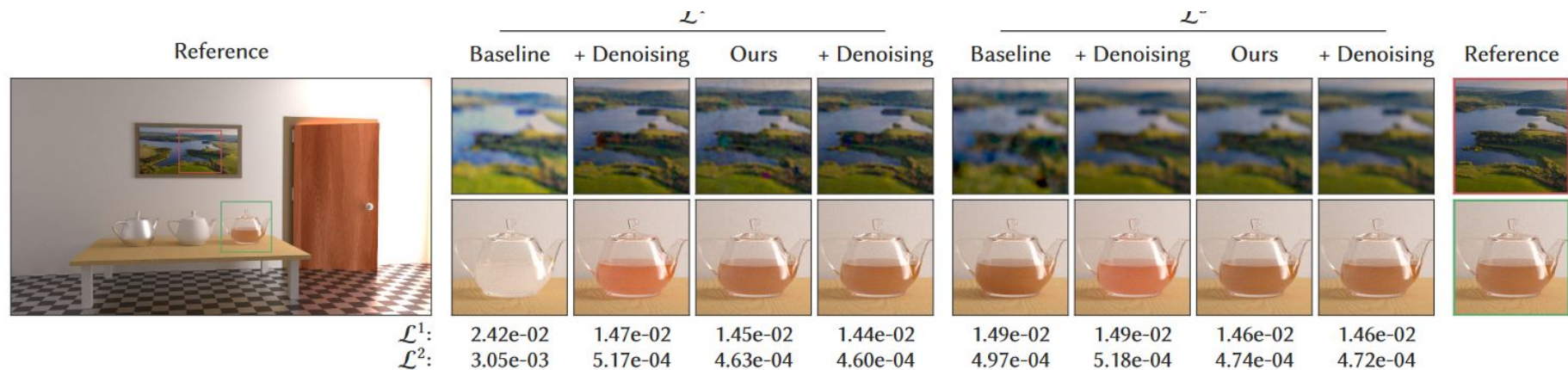
It x

Results

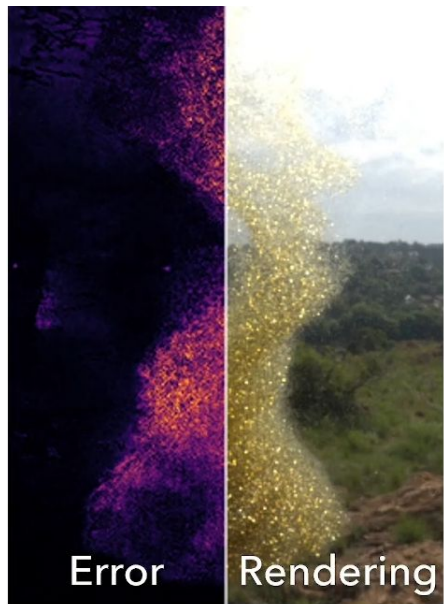


Clean rendering of optimized parameters

Results



Results



16 spp
17min45



128 spp
50min42



16 spp + CV
21min23

FIN