Supplemental Material: Adaptive Ray Marching for Rendering Gaussian Process Implicit Surfaces

ZHIQIAN ZHOU, University of California, Irvine, USA
DARIO SEYB, Valve, USA
SHUANG ZHAO, University of Illinois Urbana-Champaign, USA

This supplemental material contains two independent parts. Section 1 derives the probabilistic bounds on 1D Gaussian processes used by our adaptive ray-marching algorithm in the main paper. Section 2 describes a mathematical tool for removing a conditioning point from our online GP sampler, which may be useful for future work on optimizing the conditioning set.

1 Deriving probabilistic bounds of GP

In this section, we first introduce the reader to preliminaries in Section 1.1 and intermediate bounds in Section 1.2. Then we derive the two probabilistic bounds used by the adaptive ray-marching rule in the main paper: a long-range bound centered at the mean in Section 1.3, and a local bound centered at the currently observed value in Section 1.4. Finally we combine these two bounds and summarize the quantities used in the main paper in Section 1.5.

This analysis uses a Dudley-style chaining argument [Dudley 1967], yielding a tighter bound for 1D domains than the standard Dudley entropy integral which is for generic index sets. We note that it suffices to discuss the supremum of a function, since $\inf f = -\sup(-f)$.

1.1 Preliminaries

Theorem 1.1 (Borell–TIS inequality). *[Borell 1975; Tsirelson et al. 1976] Let $\{X_\theta\}_{\theta\in T}$ be a zero-mean Gaussian process with $X^* := \mathbb{E}\sup_{\theta\in T} X_\theta < \infty$ and $\sigma_T^2 := \sup_{\theta\in T}\operatorname{Var}[X_\theta] < \infty$. Then for all $u > 0$,*

$$\mathbb{P}\left(\sup_{\theta\in T} X_\theta > X^* + u\right) \le \exp\left(-\frac{u^2}{2\sigma_T^2}\right).$$

Lemma 1.2 (Expected max of Gaussians). *Let $(Z_1,\dots,Z_m)$ be a zero-mean (possibly correlated) Gaussian vector with $\max_i \operatorname{Var}[Z_i] \le \sigma^2$. Then*

$$\mathbb{E}\max_{1\le i\le m} Z_i \;\le\; \sigma\sqrt{2\log m}.$$

Proof. For any $\lambda > 0$,

$$\begin{aligned}\mathbb{E}\max_i Z_i &= \frac{1}{\lambda}\mathbb{E}\log\exp(\lambda\max_i Z_i)\\ &\le \frac{1}{\lambda}\log\mathbb{E}\exp(\lambda\max_i Z_i) \le \frac{1}{\lambda}\log\sum_{i=1}^m \mathbb{E}e^{\lambda Z_i}.\end{aligned}$$

Since $Z_i \sim \mathcal{N}(0,\operatorname{Var}Z_i)$, $\mathbb{E}e^{\lambda Z_i} \le \exp(\lambda^2\sigma^2/2)$. Thus $\mathbb{E}\max_i Z_i \le \frac{\log m}{\lambda} + \frac{\lambda\sigma^2}{2}$. Choose $\lambda = \sqrt{2\log m}/\sigma$. □

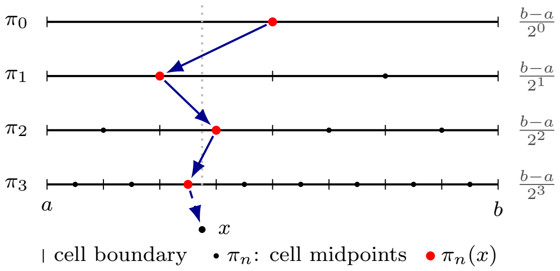

Fig. 1. Multi-resolution grid. At level n, the interval $[a,b]$ is divided into 2^n cells; π_n is the set of cell midpoints (black dots). Given any $x \in [a,b]$, $\pi_n(x) \in \pi_n$ (red) is the midpoint of the cell at level n containing x. The arrows trace the chain $\pi_0(x) \to \pi_1(x) \to \pi_2(x) \to \pi_3(x) \dashrightarrow x$; by sample-path continuity, $g(\pi_n(x)) \to g(x)$ along this chain, giving $g(x) = g(\pi_0(x)) + \sum_{n\ge1}(g(\pi_n(x)) - g(\pi_{n-1}(x)))$. The geometric decay of cell width is dominant, yielding a convergent bound on $\sup_{[a,b]} g$.

1.2 Expected supremum of zero-mean 1D GP

This subsection establishes the bound of the expected supremum of a zero-mean Gaussian process on an interval, in terms of the interval length and the kernel's smoothness. The idea is to approximate the process on a sequence of grid points (Figure 1) and bound its supremum by summing the maximum jump between successively finer grids; more points at finer levels are outweighed by geometrically smaller jumps. The derivation follows; readers may otherwise skip to Section 1.3.

Let $g \sim \mathcal{GP}(0,k)$ on $[a,b]$ and assume k is C^2 with uniformly bounded mixed derivative

$$M := \sup_{x,y\in[a,b]} \left|k_{X,Y}(x,y)\right| < \infty.$$

where $k_{X,Y}(x,y) := \frac{\partial^2}{\partial x\,\partial y}k(x,y)$.

Define the canonical metric $d(s,t) := \sqrt{\mathbb{E}[(g(s)-g(t))^2]}$. For $s < t$,

$$\begin{aligned}d(s,t)^2 &= \mathbb{E}[(g(s)-g(t))^2]\\ &= k(s,s) + k(t,t) - 2k(s,t)\\ &= \int_s^t\int_s^t k_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y\\ &\le M(s-t)^2\end{aligned}$$

Hence $d(s,t) \le \sqrt{M}\,|s-t|$.

For $n \ge 0$, define a multi-resolution grid (illustrated in Figure 1)

$$\pi_n := \left\{a + \left(i+\tfrac{1}{2}\right)\frac{b-a}{2^n} \;:\; i = 0,\dots,2^n-1\right\},$$

Authors' Contact Information: Zhiqian Zhou, zhou.zq@uci.edu, University of California, Irvine, USA; Dario Seyb, dario.seyb@gmail.com, Valve, USA; Shuang Zhao, shzhao@illinois.edu, University of Illinois Urbana-Champaign, USA.

and let $\pi_n(x)$ be the midpoint of the grid cell containing x at level n. Then $|\pi_n(x) - \pi_{n-1}(x)| = \frac{b-a}{2^{n+1}}$ and thus

$$d(\pi_n(x), \pi_{n-1}(x)) \leq \sqrt{M}\,\frac{b-a}{2^{n+1}}.$$

By sample-path continuity, $g(x) = \lim_{n\to\infty} g(\pi_n(x))$ and therefore

$$g(x) = g(\pi_0(x)) + \sum_{n=1}^{\infty} \big(g(\pi_n(x)) - g(\pi_{n-1}(x))\big),$$

We first show the absolute convergence of its expectation. For fixed x, let $\Delta_n(x) := g(\pi_n(x)) - g(\pi_{n-1}(x))$. Then $\Delta_n(x)$ is zero-mean Gaussian with

$$\mathrm{Var}[\Delta_n(x)] = d(\pi_n(x), \pi_{n-1}(x))^2 \leq M\Big(\frac{b-a}{2^{n+1}}\Big)^2.$$

Hence

$$\mathbb{E}|\Delta_n(x)| = \sqrt{\frac{2}{\pi}}\,\sqrt{\mathrm{Var}[\Delta_n(x)]} \leq \sqrt{\frac{2M}{\pi}}\,\frac{b-a}{2^{n+1}}.$$

Therefore

$$\mathbb{E}\sum_{n=1}^{\infty} |\Delta_n(x)| = \sum_{n=1}^{\infty} \mathbb{E}|\Delta_n(x)| \leq \sqrt{\frac{2M}{\pi}}\,(b-a)\sum_{n=1}^{\infty} 2^{-(n+1)} < \infty,$$

so the series is absolutely convergent in L^1 (uniformly in x).

Moreover, the grid approximation $g(\pi_n(x))$ uniformly converges to $g(x)$ in mean square, which justifies the telescoping decomposition in the chaining bound below. Using $\sup_x \sum_n a_n(x) \leq \sum_n \sup_x a_n(x)$ and $\mathbb{E}g(\pi_0(x)) = 0$, and Fubini's theorem to interchange expectation with the infinite sum (justified by the L^1 absolute convergence established above),

$$\begin{aligned}\mathbb{E}\sup_{x\in[a,b]} g(x) &\leq \sum_{n=1}^{\infty} \mathbb{E}\sup_{x\in[a,b]} \big(g(\pi_n(x)) - g(\pi_{n-1}(x))\big)\\ &= \sum_{n=1}^{\infty} \mathbb{E}\max_{u\in\pi_n} \big(g(u) - g(\pi_{n-1}(u))\big).\end{aligned}$$

For each $u \in \pi_n$, the increment is zero-mean Gaussian with $\mathrm{Var}[g(u) - g(\pi_{n-1}(u))] = d(u, \pi_{n-1}(u))^2 \leq M\big(\frac{b-a}{2^{n+1}}\big)^2$. Applying Theorem 1.2 with $m = |\pi_n| = 2^n$ yields

$$\mathbb{E}\sup_{x\in[a,b]} g(x) \leq \sum_{n=1}^{\infty}\left(\sqrt{M}\,\frac{b-a}{2^{n+1}}\right)\sqrt{2\log(2^n)} = C_1\sqrt{M}\,(b-a),$$

where

$$C_1 := \sum_{n=1}^{\infty} 2^{-(n+1)}\sqrt{2\log(2^n)} \approx 0.793135.$$

Discussion. We may obtain a tighter long-range bound via a uniformly bounded kernel k. Assume additionally $\sigma^2 := \sup_{x,y} k(x,y) < \infty$. Let $m \in \mathbb{N}_+$ and set the coarse net π_0 to contain m evenly spaced midpoints of $[a,b]$. Repeating the chaining argument gives a tighter bound.

1.3 Bound for GP with linear drift

Write $f(x) = \mu(x) + g(x)$ with $g \sim \mathcal{GP}(0,k)$. Assume

$$\sigma^2 := \sup_x k(x,x), \quad M := \sup_{x,y}\big|k_{X,Y}(x,y)\big|, \quad \Lambda_\mu := \sup_{x\neq y}\frac{|\mu(x)-\mu(y)|}{|x-y|}.$$

The long-range bound of GP is mostly driven by Λ_μ. We illustrate how a negative drift controls the supremum over an unbounded domain. For simplicity and without loss of generality we assume $\Lambda_\mu = 1$ in the following. Other values follow by rescaling the x-axis.

Let $f(x) = -x + g(x)$ on $[0,\infty)$. We would like to bound $\sup f$. We approach this problem by bounding a linear drift $[0,\infty)$ with a step function. For any $\delta > 0$, define $t_i := i\delta$, and $\alpha := C_1\sqrt{M}$.

If $\sup_{x\in[t_i,t_{i+1}]} f(x) > u$, then for some $x \in [t_i, t_{i+1}]$ we have $g(x) > u + x \geq u + t_i = u + i\delta$, hence

$$\mathbb{P}\left(\sup_{x\geq 0} f(x) > u\right) \leq \sum_{i=0}^{\infty}\mathbb{P}\left(\sup_{x\in[t_i,t_{i+1}]} g(x) > u + i\delta\right).$$

By Borell–TIS on $[t_i, t_{i+1}]$ and $\mathbb{E}\sup g \leq \alpha\delta$ (result of Section 1.2),

$$\mathbb{P}\left(\sup_{x\in[t_i,t_{i+1}]} g(x) > u + i\delta\right) \leq \exp\left(-\frac{(u-\alpha\delta+i\delta)^2}{2\sigma^2}\right).$$

We can rewrite the sum as an integral of a step function, each term being a rectangle of width δ and height $\frac{1}{\delta}\exp\big(-(u-\alpha\delta+i\delta)^2/(2\sigma^2)\big)$. Since the terms decrease in i, the step function lies below a continuous function:

$$\begin{aligned}\mathbb{P}\left(\sup_{x>0} f(x) > u\right) &\leq \frac{1}{\delta}\int_{u-\alpha\delta-\delta}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2}\right)\mathrm{d}x\\ &= \frac{\sigma}{\delta}\sqrt{\frac{\pi}{2}}\,\mathrm{erfc}\left(\frac{u-\alpha\delta-\delta}{\sqrt{2}\sigma}\right)\end{aligned}$$

Choosing $\delta = \kappa\sigma$ and solving for a $1-\eta$ bound gives

$$u \;=\; \sigma\left(\kappa(\alpha+1) + \sqrt{2}\,\mathrm{erfc}^{-1}\!\left(\kappa\eta\sqrt{\tfrac{2}{\pi}}\right)\right), \qquad 0 < \kappa\eta\sqrt{\tfrac{2}{\pi}} < 2.$$

which can be minimized over κ numerically. Choosing $\kappa = 1/(\alpha+1)$ gives the asymptotic bound $u/\sigma = O\big(\sqrt{\log(\alpha/\eta)}\big)$.

Discussion. The result can be applied to obtain a conservative bound for any mean function with corresponding Λ_μ over a finite look-ahead range. Fix a reference point t_i and consider $x \in [t_i, t_{i+1}]$. By Lipschitzness, $\mu(x) \leq \mu(t_i) + \Lambda_\mu(t_{i+1}-t_i) - \Lambda_\mu(t_{i+1}-x)$. Viewing the mean as a linear drift retro-propagated from t_{i+1} yields an upper bound for f on $[t_i, t_{i+1}]$. Again, applying the same result for $-f$ would yield a lower bound for f.

1.4 Local bound with known point value

With the same assumptions in Section 1.3, define $h(x) := g(x) - g(0)$ on $[0, x_0]$. Then h is zero-mean Gaussian and its canonical metric

$$\begin{aligned}d_h(s,t) &= \sqrt{\mathbb{E}[(h(s)-h(t))^2]}\\ &= \sqrt{\mathbb{E}[(g(s)-g(t))^2]} \leq \sqrt{M}|s-t|.\end{aligned}$$

Therefore by applying the result of Section 1.2

$$\mathbb{E}\sup_{x\in[0,x_0]} h(x) \leq C_1\sqrt{M}\,x_0.$$

Also $\mathrm{Var}[h(x)] = d(0,x)^2 \le Mx^2 \le Mx_0^2$, so Borell–TIS gives

$$\mathbb{P}\left(\sup_{x\in[0,x_0]} h(x) > \mathbb{E}\sup_{x\in[0,x_0]} h(x) + u\right) \le \exp\left(-\frac{u^2}{2Mx_0^2}\right).$$

Hence, for any $\eta \in (0,1)$,

$$\sup_{x\in[0,x_0]} h(x) \le \left(C_1 + \sqrt{-2\log\eta}\right)\sqrt{M}\,x_0 \qquad \text{with probability } \ge 1-\eta.$$

Note that this result isn't conditioned on $g(0)$.

1.5 Summary

We summarize the derived results cited in the main paper. Note that when two high-probability bounds are used simultaneously, their failure probability should be no greater than $\eta_1 + \eta_2$ where η_1, η_2 are the failure probabilities of each bound.

Assuming without loss of generality that $f(t_i) > 0$,

$$\begin{aligned}\inf_{[t_i,t_{i+1}]} f &\ge \mu(t_i) - \Lambda_\mu\Delta_i - Q_\eta\sigma,\\ \inf_{[t_i,t_{i+1}]} f &\ge f(t_i) - \left(\Lambda_\mu + C_\eta\sqrt{M}\right)\Delta_i.\end{aligned} \tag{1}$$

each with a probability no less than $(1-\eta)$. Here,

$$\begin{aligned}\Delta_i &:= t_{i+1} - t_i,\\ C_1 &\approx 0.793135,\\ Q_\eta &:= \min_{0<\kappa<\sqrt{2\pi}/\eta} \kappa\left(C_1\sqrt{M}/\Lambda_\mu + 1\right) + \sqrt{2}\ \mathrm{erfc}^{-1}\left(\kappa\eta\sqrt{\tfrac{2}{\pi}}\right),\\ C_\eta &:= C_1 + \sqrt{-2\log\eta}.\end{aligned} \tag{2}$$

2 Pruning conditioning points in online GP Sampler

The online GP sampler in the main paper accumulates conditioning points along the ray. Although our method already drastically reduces the number of samples, the conditioning set X can still grow as the ray advances. We therefore provide an option to cap the number of concurrently retained correlated samples, by keeping at most $n_{\max}$ conditioning points on a rolling basis, and evicting old ones when necessary. Note that this approximation makes the sampling no longer exact, and this technique isn't used in the main results but only for readers' reference.

This option is reminiscent of the Renewal+ model [Seyb et al. 2024], should the reader be familiar with it, in that it preserves local correlation near the current search interval while discarding distant ones. However, unlike Renewal+ which only retains a single state at fixed ray segment boundaries, our option retains $n_{\max}$ conditioning points that evolve continuously as the ray advances.

Specifically, after a conditioning point is added to X , if $|X| > n_{\max}$, we evict the i-th conditioning point from X. For a kernel that decays with distance, it is natural to take $i = 1$. The updated covariance matrix becomes the principal submatrix $k(X \setminus x_i, X \setminus x_i)$. We update the Cholesky decomposition of $k(X, X)$ accordingly: let

$$L = \begin{bmatrix} L_{11} & & \\ l_{21}^\top & a_i & \\ L_{31} & l_{32} & L_{33} \end{bmatrix} \tag{3}$$

be the original Cholesky decomposition of $k(X, X)$, where $(l_{21}^\top\ a_i)$ and $(a_i\ l_{32})^\top$ are the i-th row and i-th column of L to be removed, respectively. Then, the updated Cholesky decomposition is

$$L' = \begin{bmatrix} L_{11} & \\ L_{31} & L'_{33} \end{bmatrix}, \tag{4}$$

where $L'_{33}L'^\top_{33} = L_{33}L_{33}^\top + l_{32}l_{32}^\top$. This operation is known as the rank-1 Cholesky update [Gill et al. 1974] and has the complexity of $O((|X| - i)^2)$.

References

Christer Borell. 1975. The brunn-minkowski inequality in gauss space. *Inventiones mathematicae* 30, 2 (1975), 207–216.

Richard M Dudley. 1967. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis* 1, 3 (1967), 290–330.

Philip E Gill, Gene H Golub, Walter Murray, and Michael A Saunders. 1974. Methods for modifying matrix factorizations. *Mathematics of computation* 28, 126 (1974), 505–535.

Dario Seyb, Eugene d'Eon, Benedikt Bitterli, and Wojciech Jarosz. 2024. From microfacets to participating media: A unified theory of light transport with stochastic geometry. *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 43, 4 (July 2024). doi:10/gt5nh9

B. S. Tsirelson, I. A. Ibragimov, and V. N. Sudakov. 1976. Norms of Gaussian sample functions. In *Proceedings of the Third Japan–USSR Symposium on Probability Theory (Lecture Notes in Mathematics, Vol. 550)*. Springer-Verlag, Berlin, 20–41. Tashkent, 1975.