Supplementary: RAGA: Real Time Ray Traced Gaussian Shadow Casting for 3DGS Avatar-Scene Interaction

Aymen Mir*[1], Riza Alp Guler[3], Jian Wang[4], Peter Wonka[5], Bing Zhou[4], and Gerard Pons-Moll[1,2]

[1] Tübingen AI Center, University of Tübingen, Germany
[2] Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
[3] Imperial College London, UK
[4] Snap Inc., USA
[5] King Abdullah University of Science and Technology (KAUST), Saudi Arabia

This supplementary material provides additional background and proofs. It is organized as follows:

- Sec. 1: Background on the 3DGS avatar model.
- Sec. 2: Supplementary proofs.
- Sec. 3: Background on 3D Gaussian Splatting.

1 Background: 3DGS Avatar Model

In Sec. 3 of the main paper, we use a pose-dependent generator $\mathcal{H}(\boldsymbol{\theta}_t)$ to produce animated 3DGS avatars and introduce an avatar proxy (Sec. 3.8) for temporally stable shadow casting. We provide background on the avatar reconstruction pipeline here for self-containedness. The avatars in our work follow the Animatable Gaussians framework [2]: a pose-dependent neural network maps SMPL body poses to 3D Gaussian primitives, which are then articulated via linear blend skinning (LBS). We summarize each stage below and explain why this pipeline motivates the avatar proxy.

1.1 Template Reconstruction

Given multi-view video of a person, a character-specific canonical template is first reconstructed. Rather than relying on the naked SMPL mesh, which cannot represent loose garments such as dresses, a canonical signed distance field (SDF) and color field are optimized from a near-A-pose frame and a mesh is extracted via Marching Cubes [2]. SMPL skinning weights are diffused from the body surface to the template surface, enabling the template to be deformed via LBS while faithfully capturing the geometry of worn clothing.

*Work partly done as intern at Snap Inc.

1.2 Pose Maps and Gaussian Maps

The key idea behind Animatable Gaussians is to parameterize 3D Gaussians as 2D maps on the template surface, enabling powerful 2D convolutional architectures to predict pose-dependent Gaussian attributes.

Pose maps. For each pose $\boldsymbol{\theta}_t$, the canonical template is deformed via LBS and rendered under front and back orthographic projections, producing two position maps $\mathcal{P}_f(\boldsymbol{\theta}_t)$ and $\mathcal{P}_b(\boldsymbol{\theta}_t)$. Each pixel stores the 3D coordinates of the corresponding deformed template vertex, encoding the current body configuration as a structured 2D image. Because orthographic front and back views cover nearly the entire body surface, this simple parameterization avoids the need for explicit UV unwrapping.

Gaussian maps. A StyleUNet f_ϕ, a U-Net augmented with StyleGAN-based modulation layers [2], maps the pose maps to *Gaussian maps*:

$$\mathcal{M}_f(\boldsymbol{\theta}_t),\ \mathcal{M}_b(\boldsymbol{\theta}_t)\ \leftarrow\ f_\phi\big(\mathcal{P}_f(\boldsymbol{\theta}_t),\ \mathcal{P}_b(\boldsymbol{\theta}_t),\ \mathcal{V}\big),$$

where $\mathcal{V}$ encodes the view direction relative to the template normal. Each pixel in a Gaussian map encodes a complete 3D Gaussian primitive: a position offset $\Delta\boldsymbol{\mu}$ relative to the template surface, per-axis scales $\mathbf{s}$, a rotation quaternion $\mathbf{q}$, opacity α, and view-dependent color $\mathbf{c}$. The collection of all foreground pixels across the front and back maps yields the full set of canonical Gaussians $\mathcal{G}_t^{H,\mathsf{C}}$.

Training. The StyleUNet is trained end-to-end from multi-view video (typically 16–48 cameras). For each training frame, the predicted canonical Gaussians are posed via LBS and rendered via standard 3DGS rasterization. Parameters are optimized with a photometric reconstruction loss (L1 + perceptual) between rendered and ground-truth images.

1.3 Linear Blend Skinning for Articulation

Canonical Gaussians are articulated into posed space using linear blend skinning, following standard practice in SMPL-based [3,4] avatar models. Each canonical Gaussian k has center $\boldsymbol{\mu}_k^{(0)} \in \mathbb{R}^3$, covariance $\boldsymbol{\Sigma}_k^{(0)}$, opacity α_k, and color $\mathbf{c}_k$, and is anchored to the template surface by nearest-neighbor association. Let $\mathcal{J}$ denote the set of skeleton joints. For Gaussian k, skinning weights $w_{kj} \in [0,1]$ with $\sum_{j\in\mathcal{J}} w_{kj} = 1$ are inherited from the associated template vertex.

Given pose parameters $\boldsymbol{\theta}_t$, the SMPL model provides per-joint rigid transformations $\mathbf{M}_j(\boldsymbol{\theta}_t) \in \mathbb{R}^{4\times 4}$. The posed Gaussian center is:

$$\boldsymbol{\mu}_k^{(t)} = \sum_{j\in\mathcal{J}} w_{kj}\,\mathbf{M}_j(\boldsymbol{\theta}_t)\,\boldsymbol{\mu}_k^{(0)}, \tag{1}$$

and the posed covariance is approximated by the blended rotation:

$$\boldsymbol{\Sigma}_k^{(t)}\ \approx\ \bar{\mathbf{R}}_k(\boldsymbol{\theta}_t)\,\boldsymbol{\Sigma}_k^{(0)}\,\bar{\mathbf{R}}_k(\boldsymbol{\theta}_t)^\top, \qquad \bar{\mathbf{R}}_k(\boldsymbol{\theta}_t) = \sum_{j\in\mathcal{J}} w_{kj}\,\mathbf{R}_j(\boldsymbol{\theta}_t), \tag{2}$$

where $\mathbf{R}_j(\boldsymbol{\theta}_t) \in \mathrm{SO}(3)$ is the rotational component of $\mathbf{M}_j(\boldsymbol{\theta}_t)$.

The full pipeline defines a differentiable mapping $\mathcal{H} : \boldsymbol{\theta}_t \mapsto \mathcal{G}_t^H$ from body pose to posed 3D Gaussians.

1.4 Motivation for the Avatar Proxy

Because the StyleUNet f_ϕ is optimized with photometric losses that enforce view-dependent appearance fidelity, its per-pixel Gaussian predictions can fluctuate in position offsets, anisotropic scales, rotations, and opacities across frames, even for nearby poses. These fluctuations are largely imperceptible during rendering, as the rasterizer integrates over many overlapping Gaussians. However, when these same Gaussians are used as shadow occluders, the fluctuations directly modulate the shadow silhouette, producing temporally unstable shadows that flicker and oscillate rapidly.

The avatar proxy (Sec. 3.8 of the main paper) addresses this by decoupling the shadow-casting geometry from the render-ready geometry. The proxy freezes the generator at a canonical reference pose, replaces anisotropic Gaussians with conservative isotropic spheres (whose silhouettes are rotation-invariant), and articulates them via position-only LBS. This eliminates the three sources of temporal instability: (i) per-frame topology changes from the generator, (ii) orientation-dependent shadow silhouettes from anisotropic primitives, and (iii) rotation artifacts from full covariance deformation.

2 Supplementary Proofs

In Sec. 3 of the main paper, we state that "all formal propositions and proofs supporting the derivations below are provided in the supplementary material." We provide them here: Lemma 1 and Propositions 1–7 with their proofs, covering the quadratic ray–Gaussian expansion (Sec. 3.2), the χ^2 intersection interval (Sec. 3.2), the closed-form line integral via the error function (Sec. 3.3), the global maximum $I_{\max}$ (Sec. 3.4), bounds on the normalized factor η (Sec. 3.4), the exponentiated opacity equivalence (Sec. 3.5), and the optical-depth mapping (Sec. 3.5).

2.1 Preliminaries and Notation

We restate the setup from the main paper for self-containedness. We consider the unnormalized anisotropic Gaussian kernel

$$g(\mathbf{x}) = \exp\Big(-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\Big), \tag{3}$$

and a ray (or segment)

$$\mathbf{x}(\tau) = \mathbf{o} + \tau \mathbf{v}, \qquad \tau \in [0, 1], \tag{4}$$

where $\mathbf{o}$ is the receiver point and $\mathbf{v} = \boldsymbol{\ell} - \mathbf{o}$ is the direction toward the light. We write $\mathbf{P} = \boldsymbol{\Sigma}^{-1} \succ 0$ for brevity and define the quadratic form along the ray:

$$q(\tau) \triangleq (\mathbf{x}(\tau) - \boldsymbol{\mu})^\top \mathbf{P}(\mathbf{x}(\tau) - \boldsymbol{\mu}). \tag{5}$$

Lemma 1 (Maximum of unnormalized Gaussian). *$g(\mathbf{x}) \leq 1$ for all $\mathbf{x}$ and $g(\boldsymbol{\mu}) = 1$.*

Proof. Since $\mathbf{P} \succ 0$, the quadratic form $(\mathbf{x} - \boldsymbol{\mu})^\top \mathbf{P}(\mathbf{x} - \boldsymbol{\mu}) \geq 0$ with equality iff $\mathbf{x} = \boldsymbol{\mu}$. Therefore the exponent is ≤ 0 and $g(\mathbf{x}) = \exp(-\frac{1}{2} \cdot \text{nonnegative}) \leq 1$, with $g(\boldsymbol{\mu}) = \exp(0) = 1$.

2.2 Proof that $q(\tau)$ is a quadratic and expressions for (A, B, C)

Proposition 1 (Quadratic expansion). *Let $\boldsymbol{\delta} = \mathbf{o} - \boldsymbol{\mu}$. Then $q(\tau)$ is a quadratic polynomial*

$$q(\tau) = A\tau^2 + 2B\tau + C, \tag{6}$$

with

$$A = \mathbf{v}^\top \mathbf{P}\mathbf{v}, \quad B = \mathbf{v}^\top \mathbf{P}\boldsymbol{\delta}, \quad C = \boldsymbol{\delta}^\top \mathbf{P}\boldsymbol{\delta}. \tag{7}$$

Proof. We have $\mathbf{x}(\tau) - \boldsymbol{\mu} = \mathbf{o} + \tau\mathbf{v} - \boldsymbol{\mu} = \boldsymbol{\delta} + \tau\mathbf{v}$. Then

$$\begin{aligned} q(\tau) &= (\boldsymbol{\delta} + \tau\mathbf{v})^\top \mathbf{P}(\boldsymbol{\delta} + \tau\mathbf{v}) \\ &= \boldsymbol{\delta}^\top \mathbf{P}\boldsymbol{\delta} + \tau\, \boldsymbol{\delta}^\top \mathbf{P}\mathbf{v} + \tau\, \mathbf{v}^\top \mathbf{P}\boldsymbol{\delta} + \tau^2\, \mathbf{v}^\top \mathbf{P}\mathbf{v}. \end{aligned} \tag{8}$$

Since $\mathbf{P}$ is symmetric, $\boldsymbol{\delta}^\top \mathbf{P}\mathbf{v} = \mathbf{v}^\top \mathbf{P}\boldsymbol{\delta}$. Grouping terms yields

$$q(\tau) = (\mathbf{v}^\top \mathbf{P}\mathbf{v})\tau^2 + 2(\mathbf{v}^\top \mathbf{P}\boldsymbol{\delta})\tau + (\boldsymbol{\delta}^\top \mathbf{P}\boldsymbol{\delta}), \tag{9}$$

which gives Eq. 7.

Correspondence with the main text. In the main paper (Sec. 3.1), the whitening transform $\mathbf{L} = \mathrm{diag}(s_x^{-1}, s_y^{-1}, s_z^{-1})\mathbf{R}^\top$ maps to Gaussian coordinates $\mathbf{y} = \mathbf{L}(\mathbf{x} - \boldsymbol{\mu})$. The quadratic coefficients $A = \|\mathbf{v}'\|_2^2$, $B = \mathbf{o}'^\top \mathbf{v}'$, $C = \|\mathbf{o}'\|_2^2$ (with $\mathbf{o}' = \mathbf{L}\boldsymbol{\delta}$ and $\mathbf{v}' = \mathbf{L}\mathbf{v}$) are identical to Eq. 7 since $\mathbf{P} = \mathbf{L}^\top \mathbf{L}$.

2.3 Proof of the χ^2 intersection interval and entry/exit times

We restrict contributions to the χ^2 level set:

$$q(\tau) \leq \chi^2. \tag{10}$$

Substituting $q(\tau) = A\tau^2 + 2B\tau + C$ gives

$$A\tau^2 + 2B\tau + (C - \chi^2) \leq 0. \tag{11}$$

Proposition 2 (Intersection on the line). *Assume $A > 0$ (true because $\mathbf{P} \succ 0$ and $\mathbf{v} \neq 0$). Define the discriminant*

$$D = 4B^2 - 4A(C - \chi^2). \tag{12}$$

If $D < 0$ then Eq. 11 has no real solutions. If $D \geq 0$ then the solution set on the infinite line is the closed interval $[\tau^-, \tau^+]$ where

$$\tau^{\pm} = \frac{-2B \pm \sqrt{D}}{2A} = -\frac{B}{A} \pm \frac{1}{2A}\sqrt{D}, \quad \text{with } \tau^- \leq \tau^+. \tag{13}$$

Proof. The quadratic polynomial $p(\tau) = A\tau^2 + 2B\tau + (C - \chi^2)$ opens upward since $A > 0$. If $D < 0$, $p(\tau)$ has no real roots and is strictly positive for all τ, so $p(\tau) \leq 0$ has no solutions. If $D \geq 0$, $p(\tau)$ has two real roots given by the quadratic formula Eq. 13. For an upward-opening parabola, $p(\tau) \leq 0$ holds exactly between the roots (including endpoints), hence the solution set is $[\tau^-, \tau^+]$.

Correspondence with the main text. The center point $\tau_c = -B/A$ and perpendicular distance $d_\perp^2 = C - B^2/A$ from the main text (Eqs. 8–11) relate to the discriminant as $D = 4A(\chi^2 - d_\perp^2) = 4A\Delta$, where $\Delta = \chi^2 - d_\perp^2$ is the quantity defined in the main text (Eq. 10). The entry/exit times $\tau_{\text{in}} = \tau_c - \sqrt{\Delta/A}$ and $\tau_{\text{out}} = \tau_c + \sqrt{\Delta/A}$ follow directly from Eq. 13.

Corollary 1 (Intersection on the segment). *For a segment domain $\tau \in [0, 1]$, the contributing interval is*

$$\tilde{\tau}_{\text{in}} = \max(0, \tau^-), \qquad \tilde{\tau}_{\text{out}} = \min(1, \tau^+), \tag{14}$$

and the Gaussian contributes iff $\tilde{\tau}_{\text{in}} < \tilde{\tau}_{\text{out}}$.

Proof. Intersect the line-solution interval $[\tau^-, \tau^+]$ with the segment domain $[0, 1]$. The intersection is empty iff $\max(0, \tau^-) \geq \min(1, \tau^+)$; otherwise it is precisely the interval Eq. 14.

2.4 Proof of the closed-form truncated line integral

We define the truncated line integral

$$I_{\text{line}} \triangleq \int_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}} \exp\Big(-\tfrac{1}{2}(A\tau^2 + 2B\tau + C)\Big)\, d\tau, \tag{15}$$

where $A > 0$.

Proposition 3 (Completing the square). *We can rewrite the quadratic as*

$$A\tau^2 + 2B\tau + C = A(\tau - \tau_c)^2 + d_\perp^2, \qquad \tau_c = -\frac{B}{A}, \qquad d_\perp^2 = C - \frac{B^2}{A}. \tag{16}$$

Proof. Expand $A(\tau-\tau_c)^2 = A(\tau^2 - 2\tau_c\tau + \tau_c^2) = A\tau^2 - 2A\tau_c\tau + A\tau_c^2$. Match the linear coefficient: $-2A\tau_c = 2B \Rightarrow \tau_c = -B/A$. Then the constant term must satisfy $A\tau_c^2 + d_\perp^2 = C$, i.e. $d_\perp^2 = C - A\tau_c^2 = C - B^2/A$.

Proposition 4 (Closed form with error function). *The integral Eq. 15 equals*

$$I_{\text{line}} = \exp\Big(-\tfrac{1}{2}d_\perp^2\Big)\sqrt{\frac{\pi}{2A}}\left[\operatorname{erf}\Big(\sqrt{\tfrac{A}{2}}(\tau-\tau_c)\Big)\right]_{\tau=\tilde{\tau}_{\text{in}}}^{\tau=\tilde{\tau}_{\text{out}}}. \tag{17}$$

Proof. Using Eq. 16,

$$\begin{aligned} I_{\text{line}} &= \int_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}} \exp\Big(-\tfrac{1}{2}\big(A(\tau-\tau_c)^2 + d_\perp^2\big)\Big)\,d\tau \\ &= \exp\Big(-\tfrac{1}{2}d_\perp^2\Big)\int_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}} \exp\Big(-\tfrac{A}{2}(\tau-\tau_c)^2\Big)\,d\tau. \end{aligned} \tag{18}$$

Let $u = \sqrt{\tfrac{A}{2}}(\tau-\tau_c)$ so that $d\tau = \sqrt{\tfrac{2}{A}}\,du$. Then

$$\int \exp(-u^2)\,du = \frac{\sqrt{\pi}}{2}\operatorname{erf}(u) + \text{const}, \tag{19}$$

and therefore

$$\begin{aligned} \int_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}} \exp\Big(-\tfrac{A}{2}(\tau-\tau_c)^2\Big)\,d\tau &= \sqrt{\tfrac{2}{A}}\int_{u_{\text{in}}}^{u_{\text{out}}} \exp(-u^2)\,du = \sqrt{\tfrac{2}{A}}\cdot\frac{\sqrt{\pi}}{2}\big[\operatorname{erf}(u)\big]_{u_{\text{in}}}^{u_{\text{out}}} \\ &= \sqrt{\frac{\pi}{2A}}\left[\operatorname{erf}\Big(\sqrt{\tfrac{A}{2}}(\tau-\tau_c)\Big)\right]_{\tilde{\tau}_{\text{in}}}^{\tilde{\tau}_{\text{out}}}. \end{aligned} \tag{20}$$

Multiplying by $\exp(-d_\perp^2/2)$ gives Eq. 17.

2.5 Proof of the global maximum line integral $I_{\max}$

We claim the maximum possible line integral *within the same* χ^2 *truncation* occurs for a ray passing through the mean and aligned with the largest standard deviation axis.

Assume the covariance is parameterized as $\boldsymbol{\Sigma} = \mathbf{R}\,\operatorname{diag}(s_x^2, s_y^2, s_z^2)\,\mathbf{R}^\top$. Define $s_{\max} = \max(s_x, s_y, s_z)$.

Step 1: Reduce to Gaussian coordinates. Let $\mathbf{L} = \operatorname{diag}(s_x^{-1}, s_y^{-1}, s_z^{-1})\,\mathbf{R}^\top$ and define $\mathbf{y} = \mathbf{L}(\mathbf{x}-\boldsymbol{\mu})$. Then $g(\mathbf{x}) = \exp(-\frac{1}{2}\|\mathbf{y}\|^2)$ and the χ^2 ellipsoid becomes $\|\mathbf{y}\|^2 \le \chi^2$.

Step 2: Center pass maximizes the exponential prefactor. Along any line in $\mathbf{y}$-space, the integrand is $\exp(-\frac{1}{2}\|\mathbf{y}(\tau)\|^2) \leq 1$ with equality only at $\mathbf{y} = 0$. For fixed direction, the maximum integral over the truncated interval is achieved when the line passes through $\mathbf{y} = 0$ (i.e. through $\boldsymbol{\mu}$ in world space), because this maximizes the integrand pointwise near the center and yields the largest possible mass under a symmetric unimodal function.

More formally, in the quadratic form $q(\tau) = A\tau^2 + 2B\tau + C$, the center pass corresponds to $B = 0$ and $C = 0$, hence $d_\perp^2 = C - B^2/A = 0$, which maximizes the multiplicative factor $\exp(-d_\perp^2/2)$ in Eq. 17.

Step 3: Alignment with largest axis maximizes the traversal length in world space. For a center pass, the truncation interval in *Gaussian coordinates* is the segment of length $2\sqrt{\chi^2}$ along the normalized direction in $\mathbf{y}$-space. Mapping back to world space scales lengths by the corresponding standard deviation along that direction. The maximum world-space length for a unit-length direction is achieved when the ray aligns with the principal axis with the largest standard deviation $s_{\max}$.

Closed form. Under a unit-speed ray parameterization where τ measures world distance, the center-pass 1D integrand reduces to

$$\exp\Big(-\frac{s^2}{2s_{\max}^2}\Big), \qquad s \in \big[-s_{\max}\sqrt{\chi^2},\ s_{\max}\sqrt{\chi^2}\big], \tag{21}$$

where s is signed distance along the ray. Thus

$$I_{\max} = \int_{-s_{\max}\sqrt{\chi^2}}^{s_{\max}\sqrt{\chi^2}} \exp\Big(-\frac{s^2}{2s_{\max}^2}\Big)\, ds. \tag{22}$$

Substitute $u = s/(\sqrt{2}s_{\max})$, $ds = \sqrt{2}s_{\max}du$, to get

$$\begin{aligned}
I_{\max} &= \sqrt{2}s_{\max} \int_{-\sqrt{\chi^2/2}}^{\sqrt{\chi^2/2}} e^{-u^2}\, du = \sqrt{2}s_{\max} \cdot \frac{\sqrt{\pi}}{2}\, [\operatorname{erf}(u)]_{-\sqrt{\chi^2/2}}^{\sqrt{\chi^2/2}} \\
&= \sqrt{2}s_{\max} \cdot \frac{\sqrt{\pi}}{2} \cdot \big(\operatorname{erf}(\sqrt{\chi^2/2}) - \operatorname{erf}(-\sqrt{\chi^2/2})\big) \\
&= \sqrt{2}s_{\max} \cdot \frac{\sqrt{\pi}}{2} \cdot 2\operatorname{erf}(\sqrt{\chi^2/2}) \\
&= \sqrt{2\pi}\, s_{\max}\ \operatorname{erf}\Big(\sqrt{\tfrac{\chi^2}{2}}\Big),
\end{aligned} \tag{23}$$

using the odd symmetry $\operatorname{erf}(-x) = -\operatorname{erf}(x)$. This proves the stated maximum Eq. 23. □

Non-unit parameterization. The main paper uses $\mathbf{x}(\tau) = \mathbf{o} + \tau\mathbf{v}$ with $\tau \in [0, 1]$ and $\mathbf{v}$ not unit-length. Then $d\tau$ corresponds to world distance $ds = \|\mathbf{v}\|\, d\tau$, so the maximum in τ-units becomes

$$I_{\max} = \frac{1}{\|\mathbf{v}\|_2}\ \sqrt{2\pi}\, s_{\max}\ \operatorname{erf}\Big(\sqrt{\tfrac{\chi^2}{2}}\Big), \tag{24}$$

matching Eq. 17 in the main text.

2.6 Justification of the normalized factor semantics

Proposition 5 (Bounds for η). *With $\eta = I_{\text{line}}/I_{\max}$ and $I_{\max} > 0$, we have $\eta \geq 0$. If additionally $I_{\text{line}} \leq I_{\max}$ (i.e., if $I_{\max}$ is a valid global upper bound for the chosen parameterization), then $\eta \in [0, 1]$.*

Proof. The integrand in Eq. 15 is nonnegative, so $I_{\text{line}} \geq 0$ and thus $\eta \geq 0$. If $I_{\text{line}} \leq I_{\max}$, then $\eta = I_{\text{line}}/I_{\max} \leq 1$. In practice we also clamp η to $[0, 1]$.

2.7 Equivalence between exponentiated attenuation and effective opacity

Proposition 6 (Equivalence of the two exponentiated forms). *Let $\alpha \in [0, 1)$ and $\eta \geq 0$. Define $\alpha_{\text{eff}} = 1 - (1-\alpha)^\eta$. Then*

$$(1 - \alpha)^\eta = 1 - \alpha_{\text{eff}}. \tag{25}$$

Therefore $T \leftarrow T(1-\alpha)^\eta$ is identical to $T \leftarrow T(1 - \alpha_{\text{eff}})$.

Proof. By definition, $\alpha_{\text{eff}} = 1 - (1-\alpha)^\eta$, rearranging gives $(1-\alpha)^\eta = 1 - \alpha_{\text{eff}}$.

Corollary 2 (Thickness additivity). *For two traversals with factors η_1, η_2, the combined transmittance multiplier is*

$$(1-\alpha)^{\eta_1}(1-\alpha)^{\eta_2} = (1-\alpha)^{\eta_1+\eta_2}. \tag{26}$$

Proof. Immediate from exponent rules.

2.8 Optical-depth mapping $\kappa(\alpha) = -\ln(1 - \alpha)$

Proposition 7 (From alpha to optical depth). *For $\alpha \in [0, 1)$, define $\kappa = -\ln(1-\alpha) \geq 0$. Then*

$$1 - \alpha = e^{-\kappa}. \tag{27}$$

Proof. Take exponentials: $e^{-\kappa} = e^{\ln(1-\alpha)} = 1 - \alpha$. Since $1 - \alpha \in (0, 1]$, $-\ln(1 - \alpha) \geq 0$.

Corollary 3 (Small-α approximation). *As $\alpha \to 0$, $\kappa = -\ln(1-\alpha) = \alpha + O(\alpha^2)$.*

Proof. Use the Taylor expansion $\ln(1-x) = -(x + \frac{x^2}{2} + \cdots)$ for $|x| < 1$.

3 Background: 3D Gaussian Splatting

For completeness, we briefly review the 3D Gaussian Splatting (3DGS) rendering framework [1] that underlies both the scene representation and the avatar model used in our work.

3DGS represents a static 3D scene with a set of Gaussian primitives $\mathcal{G}$, each parameterized by its center $\boldsymbol{\mu} \in \mathbb{R}^3$, covariance $\boldsymbol{\Sigma} \in \mathbb{R}^{3\times 3}$, opacity $\alpha \in (0,1)$, and view-dependent color $\mathbf{c}$. To ensure positive semi-definiteness, the covariance is decomposed into rotation $\mathbf{R}$ and scale $\mathbf{s}$ components:

$$\boldsymbol{\Sigma} = \mathbf{R}\ \mathrm{diag}(s_x^2,\ s_y^2,\ s_z^2)\ \mathbf{R}^\top, \tag{28}$$

where a quaternion $\mathbf{q} \in \mathbb{R}^4$ represents the rotation and a scaling vector $\mathbf{s} \in \mathbb{R}^3$ parameterizes the diagonal scale matrix. Each Gaussian k is thus specified by the tuple $(\boldsymbol{\mu}_k, \mathbf{q}_k, \alpha_k, \mathbf{s}_k, \mathbf{c}_k)$.

Rendering. The 3D Gaussians are projected to the 2D image plane and alpha-blended. Under camera projection, the 2D covariance in screen space [5] is $\boldsymbol{\Sigma}_k^{2D} = (\mathbf{J}\,\boldsymbol{\Sigma}_k\,\mathbf{J}^\top)_{1:2,1:2}$, where $\mathbf{J}$ is the Jacobian of the affine approximation of the projective transformation. The pixel color $\hat{\mathbf{I}}(\mathbf{u})$ is computed by blending Gaussian splats that overlap at pixel $\mathbf{u}$, sorted by depth:

$$\hat{\mathbf{I}}(\mathbf{u}) = \sum\nolimits_i \Big(\hat{\alpha}_i(\mathbf{u}) \prod\nolimits_{j=1}^{i-1} \big(1 - \hat{\alpha}_j(\mathbf{u})\big)\Big)\mathbf{c}_i, \tag{29}$$

where $\hat{\alpha}_i(\mathbf{u}) = \alpha_i\, g_i(\mathbf{u})$ is the learned opacity α_i weighted by the 2D Gaussian kernel $g_i(\mathbf{u})$ evaluated at pixel $\mathbf{u}$.

Optimization. The parameters of the individual 3D Gaussians are optimized by comparing rendered pixels against ground-truth images via a photometric loss. During optimization, 3DGS adaptively controls the number of primitives through periodic densification and pruning.

References

1. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics **42**(4) (2023)
2. Li, Z., Zheng, Z., Wang, L., Liu, Y.: Animatable gaussians: Learning pose-dependent gaussian maps for high-fidelity human avatar modeling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2024)
3. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: A skinned multi-person linear model. ACM Transactions on Graphics **34**(6) (2015)
4. Pavlakos, G., Choutas, V., Ghorbani, N., Bolkart, T., Osman, A.A.A., Tzionas, D., Black, M.J.: Expressive body capture: 3D hands, face, and body from a single image. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
5. Zwicker, M., Pfister, H., van Baar, J., Gross, M.: Surface splatting. In: ACM SIGGRAPH Conference Proceedings (2001)