1 PROOFS

In the following we provide mathematical proof for various statements that are used within the paper.

1.1 Case Distinction for Shading

A convenient aspect of our non-linear representation of the moments is that \( y_1 \) and \( y_2 \) directly characterize the three different cases that occur in moment shadow mapping. In the following we assume \( y_1 < y_2 \) and \( z_1 < z_2 \). Then we claim that the shadow intensity is zero for \( z_f \in (-\infty, y_1) \), \( w_1 \) for \( z_f \in (y_1, y_2) \) and \( w_1 + w_2 = 1 - w_0 \) for \( z_f \in (y_2, \infty) \). Here is a formalization of this statement.

**Proposition 1.** For moments \( b \in \mathbb{R}^3 \) with \( b_0 = 1 \) and positive definite Hankel matrix \( B(b) \) let \( y_1, y_2 \in \mathbb{R} \) be the outputs of Algorithm 2. Let

\[
z_1, z_2 : \mathbb{R} \setminus \{y_1, y_2\} \rightarrow \mathbb{R}
\]

describe the values computed by Algorithm 1 where \( z_1(z_f) < z_2(z_f) \) for all \( z_f \in \mathbb{R} \setminus \{y_1, y_2\} \). Let

\[
N : \mathbb{R} \setminus \{y_1, y_2\} \rightarrow \{0, 1, 2\}
\]

\[
z_f \mapsto \left\{ \left[ z \in \{z_1(z_f), z_2(z_f)\} \mid z < z_f \right] \right\}
\]

be a map counting the weights that contribute to the shadow intensity. Then

\[
N(z_f) = \begin{cases} 
0 & \text{if } z_f \leq y_1, \\
1 & \text{if } y_1 < z_f \leq y_2, \\
2 & \text{otherwise}.
\end{cases}
\]

Proof. As a first step, we prove that \( z_1(z_f) \) and \( z_2(z_f) \) are well-defined. According to the correctness proof of Algorithm 1 [Peters and Klein 2015, Proposition 3], these quantities are ill-defined if and only if \( q_2 = 0 \). Thus, we have to prove that \( q_2 = 0 \) if and only if \( z_f \in \{y_1, y_2\} \). Let \( e_2 := (0, 0, 1)^T \) denote a canonical basis vector.

Then by definition of \( q \):

\[
q_2 = 0 \iff (\det b(b) \cdot e_2^T \cdot B^{-1}(b)) \cdot b(z_f) = 0
\]

\[
\iff (b_2 - b_2^2) \cdot z_f^2 + (b_1 - b_2 \cdot b_1) \cdot z_f + b_1 \cdot b_3 - b_2^2 = 0
\]

This quadratic polynomial agrees with the polynomial in Algorithm 2 which has the roots \( y_1, y_2 \).

The functions \( z_1(z_f) \) and \( z_2(z_f) \) are compositions of continuous functions with no singularities besides \( y_1 \) and \( y_2 \). Therefore, they are continuous on \( \mathbb{R} \setminus \{y_1, y_2\} \). Except for their order, they are fully characterized by the fact that the following matrix is an invertible diagonal matrix for all \( z_f \in \mathbb{R} \setminus \{y_1, y_2\} \) [Peters and Klein 2015, Proposition 10]:

\[
\begin{pmatrix}
1 & 1 & 1 \\
z_f & z_1(z_f) & z_2(z_f) \\
z_1^2 & z_1^2(z_f) & z_2^2(z_f)
\end{pmatrix}^T \cdot B^{-1}(b) \cdot \begin{pmatrix}
1 & 1 & 1 \\
z_f & z_1(z_f) & z_2(z_f) \\
z_1^2 & z_1^2(z_f) & z_2^2(z_f)
\end{pmatrix}
\]

It follows that the set \( \{z_1(z_f), z_2(z_f)\} \) always has cardinality three because otherwise two matrices in this product would not be invertible. Furthermore, this set does not change if \( z_f \) is replaced by any other element of \( \{z_1(z_f), z_2(z_f)\} \) because that only permutes rows and columns of the diagonal matrix in a symmetric fashion.

Suppose \( N \) has a discontinuity at \( z \in \mathbb{R} \setminus \{y_1, y_2\} \). Then either the inequality \( z_1(z_f) < z_f \) or \( z_2(z_f) < z_f \) changes at \( z_f = z \). Since both functions are continuous at \( z \) this implies either \( z_1(z) = z \) or \( z(z) = z \). Contradiction.

Finally, we note that a minimal choice of \( z_0 \in \{z_1(z_f), z_2(z_f)\} \) has to lead to \( N(z_0) = 0 \), while the largest choice leads to \( N(z_0) = 2 \) and the choice in between leads to \( N(z_0) = 1 \). Considering that \( N \) is continuous on \( \mathbb{R} \setminus \{y_1, y_2\} \), this completes our proof.

1.2 Bounds on \( y_1, y_2 \)

Our quantization scheme exploits \( y_1, y_2 \in [-1, 1] \). The proof of this statement is a simple consequence of Proposition 1.

**Proposition 2.** Let \( Z \) be a distribution on \([-1, 1]\), let \( b = \mathcal{E}_Z(b) \) and assume that \( b_2 - b_2^2 > 0 \). Let \( y_1, y_2 \) be the outputs of Algorithm 2 with \( y_1 < y_2 \). Then \( y_1, y_2 \in [-1, 1] \).

Proof. Suppose \( y_1 < -1 \). Consider \( z_f \in \mathbb{R} \) with \( y_1 < z_f < -1 \) and \( z_f < y_2 \). By Proposition 1, the shadow intensity for \( z_f \) is given by

\[
\omega_1 = \frac{1}{(1, z_1, z_1^2)^T \cdot B^{-1}(b) \cdot (1, z_1, z_1^2)} > 0.
\]

Obviously, this cannot be a lower bound to \( Z(z < z_f) = 0 \). Contradiction.

Proving \( y_2 \leq 1 \) works analogously by considering \( z_f \in \mathbb{R} \) with \( 1 < z_f < y_2 \) and \( y_1 < z_f \) and optimal upper bounds.
1.3 Bounds on $\xi_4$

Another inequality used by our quantization scheme is $\xi_4 \leq 0.25$.

Proposition 3. Let $Z$ be a distribution on $[-1, 1]$, let $b = E_Z(b)$ and assume that $b_2 - b_4^2 > 0$. Let $\xi_4$ be the output of Algorithm 2. Then $\xi_4 \leq 0.25$ and this bound is sharp.

Proof. For $z \in [-1, 1]$ we know $z^4 \leq z^2$ and thus

$$b_4 = E_Z(z^4) \leq E_Z(z^2) = b_2.$$ 

Now we consider the definition of $\xi_4$ from Algorithm 2:

$$\xi_4 = b_4 - b_2^2 - \frac{(b_3 - b_1 \cdot b_2)^2}{b_2 - b_4^2} \leq b_4 - b_2^2 \leq b_2 - b_2^2.$$ 

Basic analysis shows that $b_2 - b_2^2$ takes its global maximum at $b_2 = 0.5$. Thus,

$$\xi_4 \leq b_2 - b_2^2 \leq 0.5 - 0.5^2 = 0.25.$$ 

To show that this upper bound is sharp, we consider the case

$$Z = 0.25 \cdot \delta_{-1} + 0.5 \cdot \delta_0 + 0.25 \cdot \delta_1.$$ 

Then

$$b = (1, 0, 0.5, 0, 0.5)$$

and therefore

$$\xi_4 = b_4 - b_2^2 - 0 = 0.5 - 0.5^2 = 0.25.$$ 

1.4 Scaling of $\xi_4$

For shading with a minimal number of operations we shift and scale the domain of depth values to achieve $y_1 = 0$ and $y_2 = 1$. This transform necessitates a change of $\xi_4$ that we derive below.

Proposition 4. Let $Z$ be a depth distribution on $\mathbb{R}$, let $b = E_Z(b)$ and assume $b_2 - b_4^2 > 0$. Let $y_1, y_2, v_4 \in \mathbb{R}$ be the outputs of Algorithm 2 for input $b$. Let

$$c := \frac{1}{y_2 - y_1}, \quad d := -c \cdot y_1$$

be the scaling and shifting needed to normalize the depth values. For $j \in \{0, 1, 2, 3, 4\}$ let

$$b'_j := E_Z((c \cdot z + d)^j)$$

denote the moments of the scaled and shifted depth distribution. Let $\xi_4'$ be the output of Algorithm 2 for input $b'$. Then

$$\xi_4' = c^4 \cdot \xi_4.$$ 

Proof. Let $S := (1 - v_2) \cdot \delta_{y_1} + v_2 \cdot \delta_{y_2}$ be the sparse distribution that reproduces the moments $b_1, b_2, b_3$ [Peters et al. 2017, Proposition 1]. Then by definition

$$\xi_4 = E_Z(z^4) - E_S(z^4).$$ 

Scaling and shifting just applies a linear transform to the moments [Peters et al. 2017, Equation (2)]. Thus, scaling and shifting both $Z$ and $S$ leads to a new pair of depth distributions where the first three moments agree. Since nothing changes about the sparsity of $S$, we obtain

$$\xi_4' = E_Z((c \cdot z + d)^4) - E_S((c \cdot z + d)^4)$$

$$= E_Z \left( \sum_{j=0}^{4} \left( \begin{array}{c} 4 \\ j \end{array} \right) c^4 \cdot z^j \cdot d^{4-j} \right) - E_S \left( \sum_{j=0}^{4} \left( \begin{array}{c} 4 \\ j \end{array} \right) c^j \cdot z^j \cdot d^{4-j} \right)$$

$$= \sum_{j=0}^{4} \left( \begin{array}{c} 4 \\ j \end{array} \right) c^j \cdot (E_Z(z^j) - E_S(z^j)) \cdot d^{4-j}.$$ 

In this sum, the terms for $j \in \{0, 1, 2, 3\}$ vanish because the corresponding moments of $S$ and $Z$ agree. What remains is

$$\xi_4' = \left( \begin{array}{c} 4 \\ 4 \end{array} \right) c^4 \cdot (E_Z(z^4) - E_S(z^4)) \cdot d^0 = c^4 \cdot \xi_4.$$

REFERENCES
