OLMo-7B + CLIP: L2 Norm Distribution
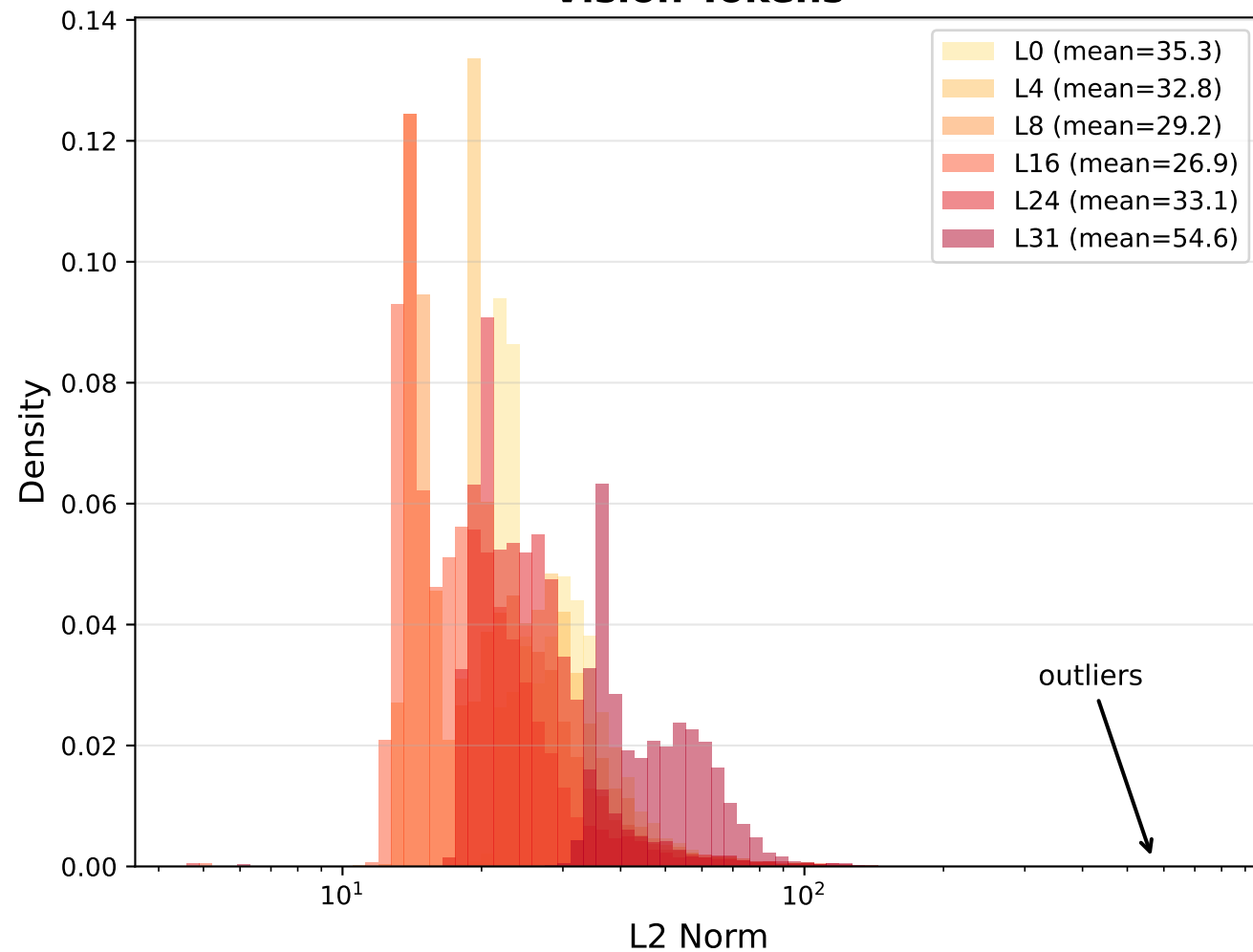
Vision Tokens

- L0 (mean=35.3)
- L4 (mean=32.8)
- L8 (mean=29.2)
- L16 (mean=26.9)
- L24 (mean=33.1)
- L31 (mean=54.6)

outliers

Text Tokens

- L0 (mean=3.6)
- L4 (mean=5.6)
- L8 (mean=8.8)
- L16 (mean=15.4)
- L24 (mean=34.6)
- L31 (mean=65.1)

outliers