

**SmolLM-360M: RoPE vs DroPE (Unconverted)**  
**Layer 0 identical = Massive values are in projection weights**  
**Later layers differ = Attention patterns affect residual stream**

