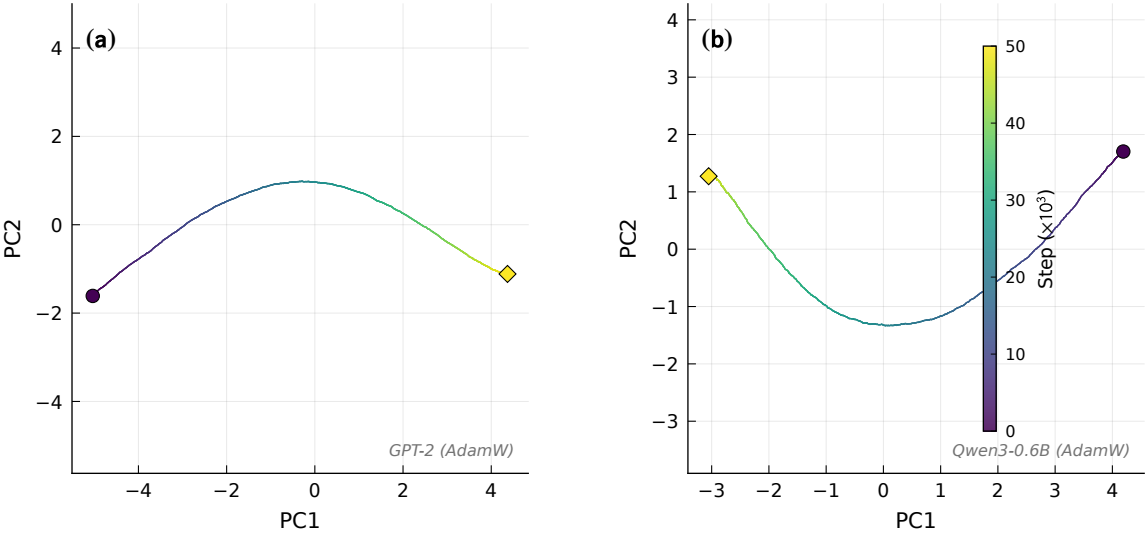
(a)
PC2
PC1
GPT-2 (AdamW)
(b)
PC2
PC1
Step (×10^3)
Qwen3-0.6B (AdamW)