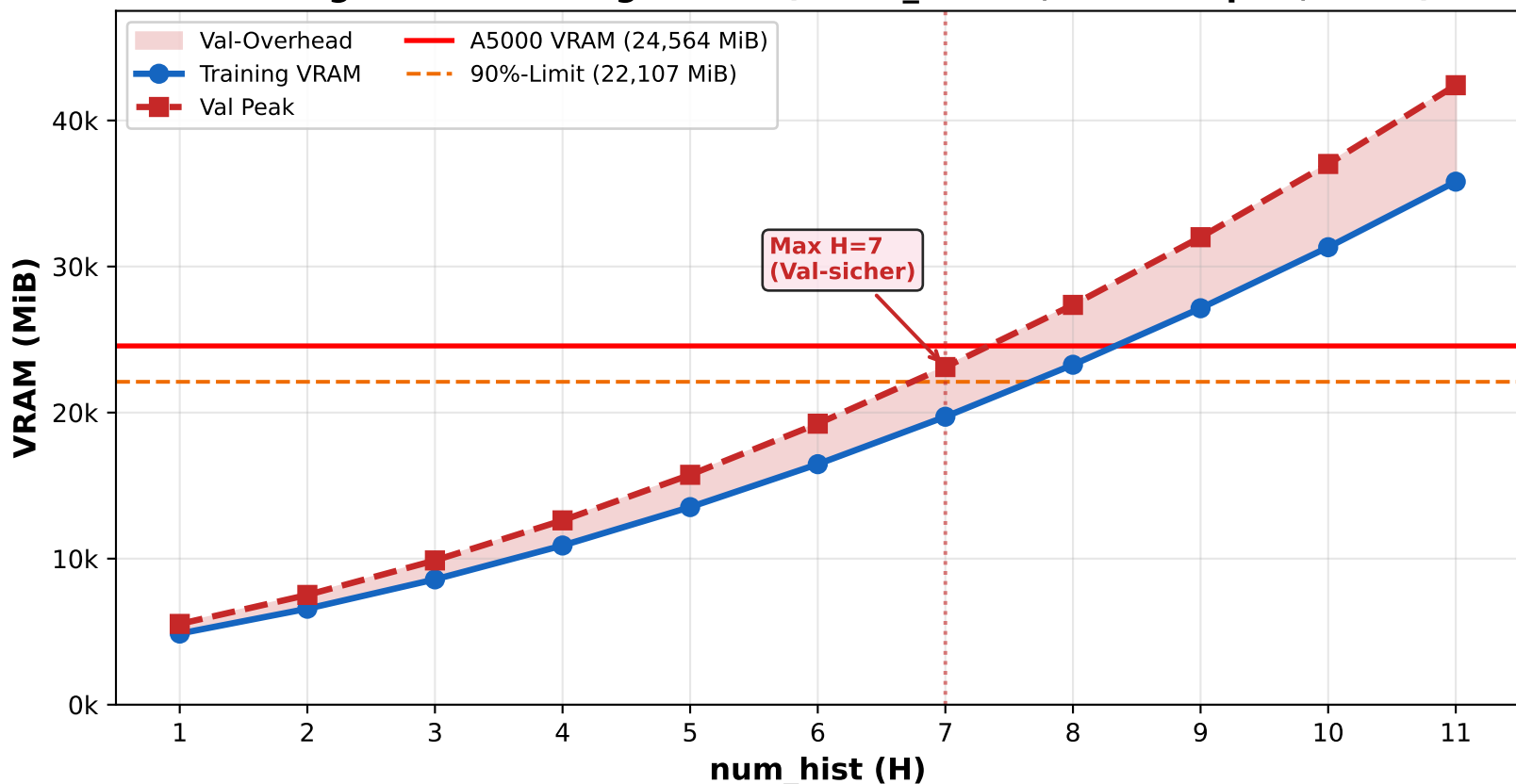
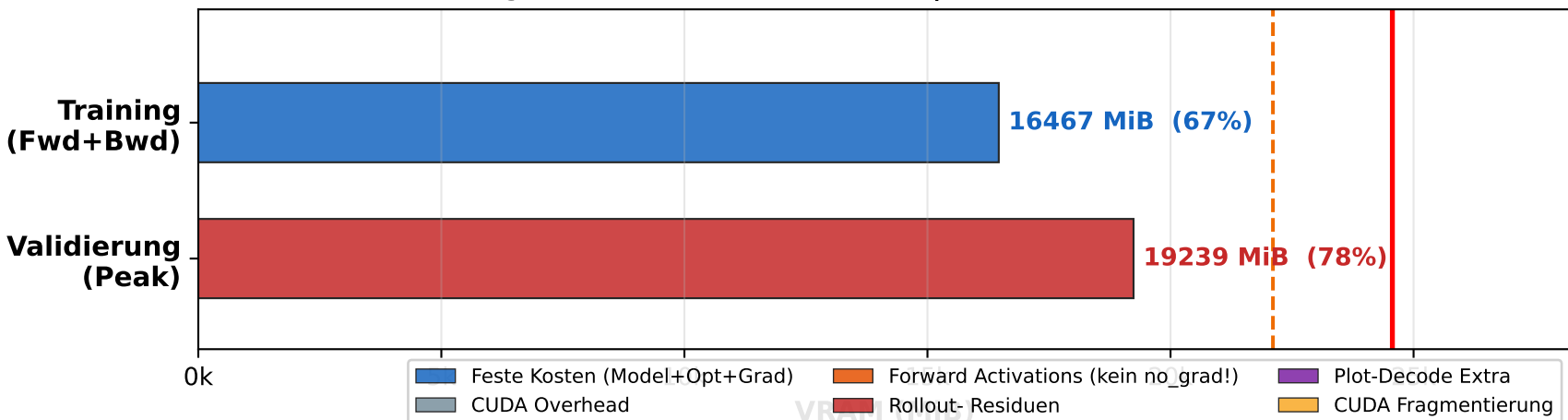


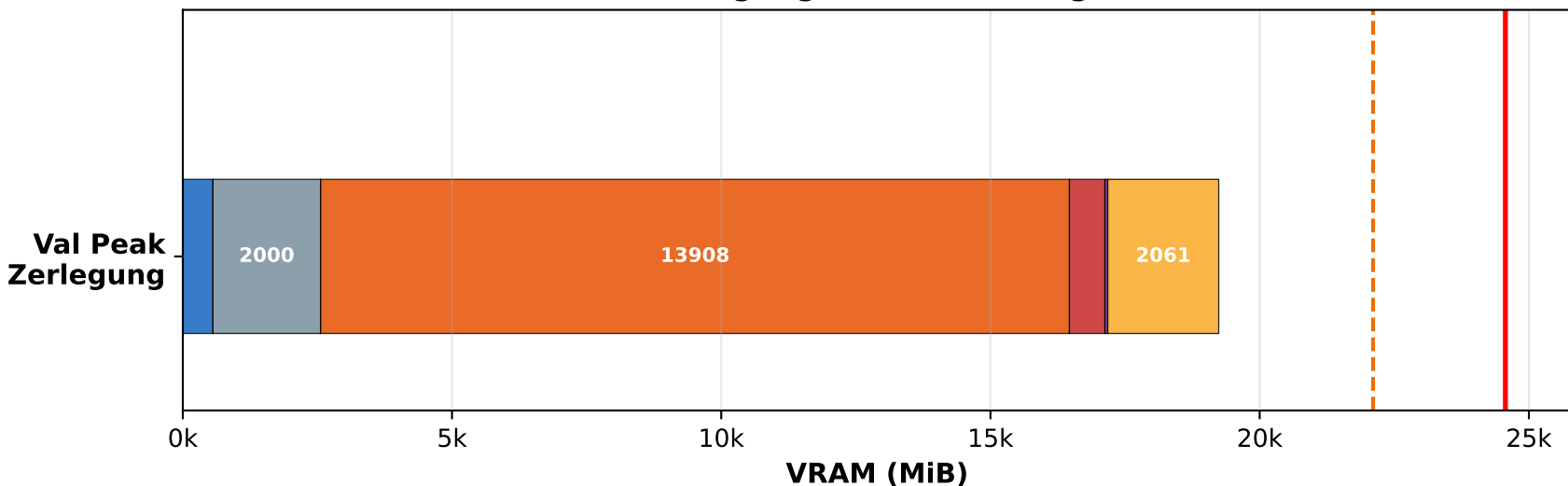
Training vs. Validierungs-Peak (batch_size=4, frameskip=2, T=22)



Aktuelle Konfiguration: bs=4, H=6, fs=2 | Val-Overhead: +2772 MiB (+17%)



VRAM-Zerlegung des Validierungs-Peaks



Ursachen des Val-Overheads:
 1. model() OHNE torch.no_grad()
 → Computation Graph wird gebaut
 2. openloop_rollout() vorher:
 → CUDA-Speicher fragmentiert
 3. Kein empty_cache() dazwischen
 → Blöcke nicht zusammenhängend