



(A) w/o all components



(B) w/o Motion-aware Diffusion Inversion



(C) w/o Motion-aware Noise Initialization



(D) w/o Probabilistic Frame Sampling



(E) Ours