

Original generation
Steering unknown latent
Random steering
Unknown latent setting
Orthogonalized model
Unknown latent
Steering known latent
Random steering
Known latent setting