Table 1: Data for Problem 1.19

Model	Steps	NLL (bits/char)(↓)
Discrete Flow (8×3 layers)	-	1.23
Argmax Coupling Flow	-	1.80
IAF / SCF	-	1.88
Multinomial Diffusion (D3PM uniform)	1000	1.72
OA-Transformer	250	1.64
D3PM-uniform	1000	1.61 $\pm$0.020
D3PM-absorbing	1000	1.45 $\pm$0.020
D3PM-absorbing	256	1.47
OA-ARDM (ours)	250	1.43 $\pm$0.001
Transformer decoder	256	1.18
Transformer XL	256	1.08
D3PM-absorbing	20	1.56 $\pm$0.001
Parallelized OA-ARDM (ours)	20	1.51 $\pm$0.001

Table 2: Data for Problem 1.19

Model	Steps	NLL (bits/dimension)(↓)
ARDM-OA	3072	2.69 $\pm$0.005
Parallel ARDM-OA	50	2.74
ARDM-Upscale 4	4×3072	2.64 $\pm$0.002
Parallel ARDM-Upscale 4	4×50	2.68
D3PM Absorbing	1000	4.40
D3PM Gaussian	1000	3.44 $\pm$0.007