

## A ORGANIZATION

This appendix is organized as follows.

- Appendix [B](#) contains further experimental results.
- Appendix [C](#) contains experimental details about datasets and hyperparameters.

## B FURTHER RESULTS

Table 3: Extended results for 3D environments – all ablations shown. Unconditional refers to DWM generating trajectories in ‘action-unconditional’ mode.

Method	CS:GO			Driving			Sample rate (Hz) ↑	Parameters (#)
	FID ↓	FVD ↓	LPIPS ↓	FID ↓	FVD ↓	LPIPS ↓		
DreamerV3	106.8	509.1	0.173	167.5	733.7	0.160	266.7	181M
IRIS ( $K = 16$ )	24.5	110.1	0.129	51.4	368.7	0.188	4.2	123M
IRIS ( $K = 64$ )	22.8	85.7	0.116	44.3	276.9	0.148	1.5	111M
DWM frame-stack, $\mathcal{T} = 20$	9.6	34.8	0.107	16.7	80.3	0.058	7.4	122M
DWM frame-stack, $\mathcal{T} = 20$ , CFG $w = 1.0$	9.7	35.1	0.108	17.6	81.5	0.059	3.7	122M
DWM frame-stack, $\mathcal{T} = 20$ , CFG $w = 2.0$	42.7	315.5	0.161	100.8	1228.6	0.198	3.7	122M
DWM frame-stack, $\mathcal{T} = 20$ , unconditional	10.1	44.3	0.120	16.7	78.7	0.058	7.4	122M
DWM frame-stack, $\mathcal{T} = 10$	14.2	60.9	0.107	23.1	100.9	0.058	14.8	122M
DWM cross-attention, $\mathcal{T} = 20$	11.6	81.4	0.125	35.2	299.9	0.119	2.5	184M
DWM cross-attention, $\mathcal{T} = 20$ , CFG $w = 1.0$	11.6	85.8	0.125	36.6	291.5	0.119	1.2	184M
DWM cross-attention, $\mathcal{T} = 20$ , CFG $w = 2.0$	13.1	99.1	0.140	107.5	1977.0	0.235	1.2	184M
DWM cross-attention, $\mathcal{T} = 20$ , unconditional	11.7	93.9	0.131	36.5	297.4	0.123	2.5	184M
DWM cross-attention, $\mathcal{T} = 10$	17.9	89.9	0.125	50.7	263.3	0.119	5.0	184M

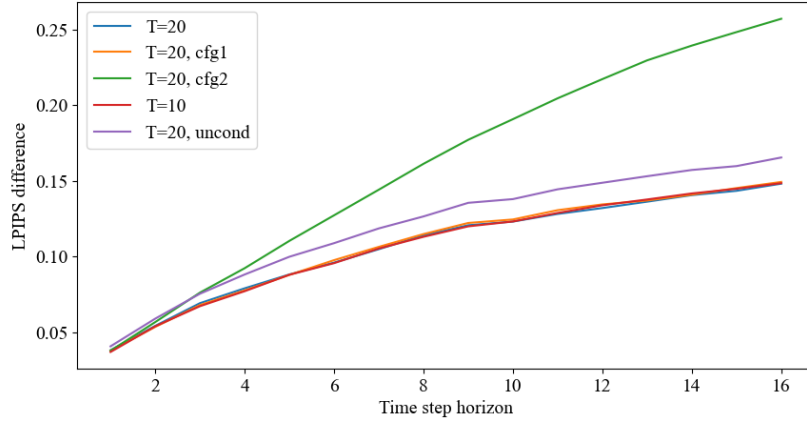


Figure 5: This figure visualizes how the LPIPS increases for increasing prediction horizon. DWM frame-stack shown on CS:GO under various sampling choices.

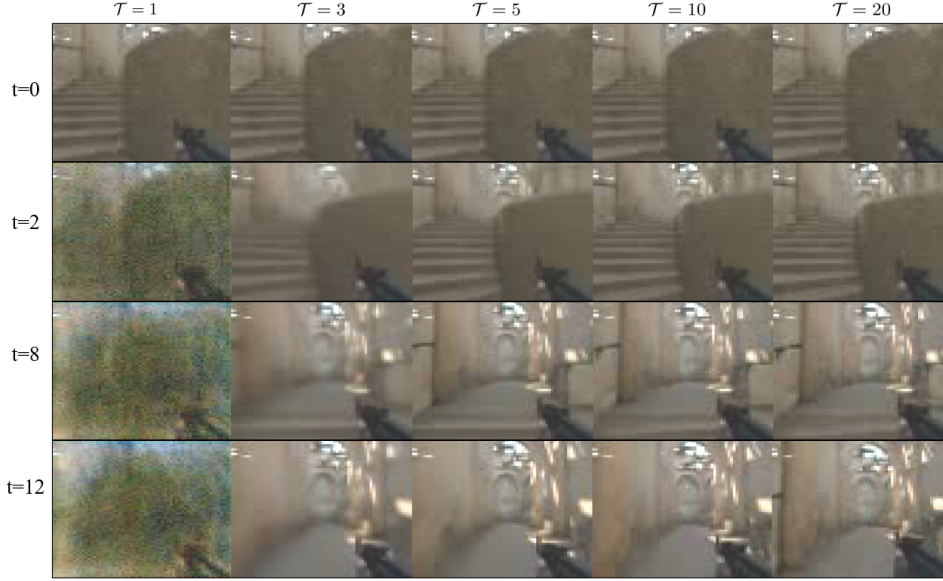


Figure 6: Effect of number of diffusion sampling steps,  $\mathcal{T}$ , on visual quality of generated observations. Diffusion world model frame-stack, CS:GO. Note the blur on the left-hand side of the generated image when  $\mathcal{T} = 3$ , and that the observations almost immediately collapse into noise when  $\mathcal{T} = 1$ . This was not the case for Atari, where very low sampling steps were successful.



Figure 7: Effect of fixed actions on sampled trajectories. Diffusion world model frame-stack, CS:GO. Conditioned on the same initial observation, we rollout the model applying differing actions. Whilst in immediate frames these have the intended effect, for longer roll-outs the observations can degenerate. For instance, it would have been very unlikely for the human demonstrator to look directly into ground in this game state, so the world model is unable to generate a plausible trajectory here, and instead snaps onto another area of the map when looking down does make sense.

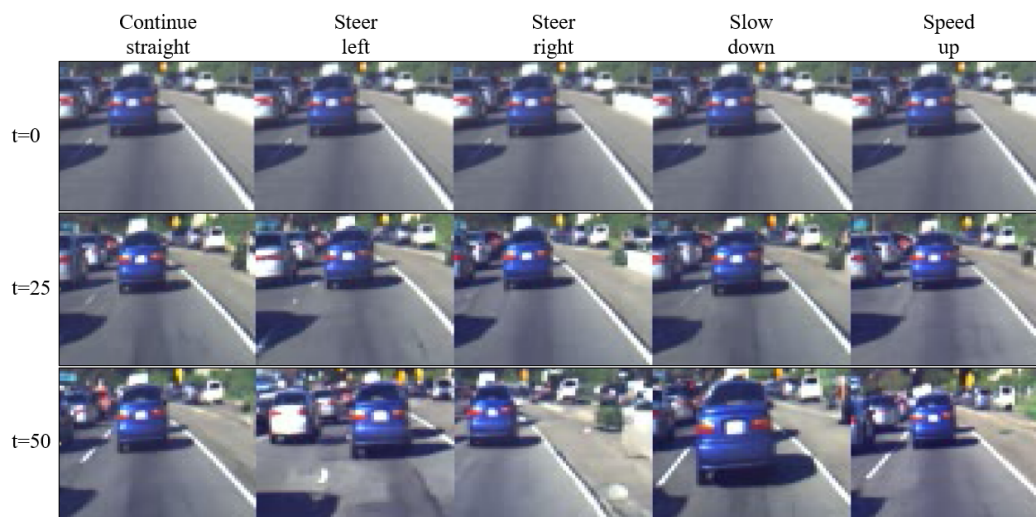


Figure 8: Effect of user-input actions on sampled trajectories. Conditioned on the same initial observations, we rollout the model applying differing actions. Note that 'Slow down' and 'Speed up' correspond to the whole traffic slowing down and speeding up.

## C EXPERIMENTAL DETAILS

### C.1 ATARI

Table 4: Hyperparameters for diffusion world model, Atari.

Hyperparameter	Value
<b>Data</b>	
Image observation dimension	$64 \times 64 \times 3$
Action space	Discrete (up to 18 actions)
<b>Model</b>	
Number previous observations ( $H$ )	4
Number previous actions	4
Prob mask actions	0.0
Prob mask observation history	0.0
Residual blocks	[2, 2, 2, 2]
U-Net channels	[64, 64, 64, 64]
Conditioning mechanism	Frame stacking
<b>Optimization</b>	
Updates	422k
Batchsize	32
Learning rate	$1e-4$
Weight decay	$1e-2$
Optimizer	AdamW
<b>Sampling</b>	
Sampler	1-step sampling
<b>Data Augmentation</b>	
Contrast	None
Brightness	None
Saturation	None
Horizontal flip probability	None
Random shift	None
<b>RI in imagination</b>	
Updates	422k
Batchsize	32
Learning rate	$1e-4$
Weight decay	$1e-2$
Optimizer	AdamW
Imagination horizon	20
Reward clipping	$\{-1, 0, 1\}$
$\gamma$	0.997
$\lambda$	0.95
Weight of entropy regularizer	0.001

## C.2 CS:GO

Table 5: Hyperparameters for diffusion world model, CS:GO.

Hyperparameter	Value
<b>Data</b>	
Image observation dimension	$64 \times 64 \times 3$
Action dimension	9
	Continuous: mouse x, mouse y
	Discrete: left click, w, a, s, d, space, r
<b>Model</b>	
Number previous observations ( $H$ )	6
Number previous actions	1
Prob mask actions	0.5
Prob mask observation history	0.5
Residual blocks	[2, 4, 4, 4]
U-Net channels	[128, 256, 512, 512]
Self-attention?	[No, Yes, Yes, Yes]
Cross-attention (when applicable)?	[No, No, Yes, Yes]
<b>Optimization</b>	
Updates	120k
Batchsize	64
Learning rate	$1e-4$
Weight decay	$1e-3$
Optimizer	AdamW
<b>Sampling</b>	
Sampler	LMS
<b>Data Augmentation</b>	
Contrast	None
Brightness	None
Saturation	None
Horizontal flip probability	None
Random shift	None

## C.3 MOTORWAY DRIVING

Table 6: Hyperparameters for diffusion world model, motorway driving.

Hyperparameter	Value
<b>Data</b>	
Image observation dimension	$64 \times 64 \times 3$
Action dimension	2
	Continuous: steer angle, accel
<b>Model</b>	
Number previous observations ( $H$ )	6
Number previous actions	1
Prob mask actions	0.2
Prob mask observation history	0.2
Residual blocks	[2, 4, 4, 4]
U-Net channels	[128, 256, 512, 512]
Self-attention?	[No, Yes, Yes, Yes]
Cross-attention (when applicable)?	[No, No, Yes, Yes]
<b>Optimization</b>	
Updates	120k
Batchsize	64
Learning rate	$1e-4$
Weight decay	$1e-3$
Optimizer	AdamW
<b>Sampling</b>	
Sampler	LMS
<b>Data Augmentation</b>	
Contrast	100% – 120%
Brightness	100% – 120%
Saturation	100% – 120%
Horizontal flip probability	0.5
Random shift	$\pm 10\%$