

342 A Model details

343 **Equations for the four-learning rate case.** In Sec. 3, for the purpose of clarity, we presented
 344 the equations for our SRWM model in the case where we only have a single learning rate β_t .
 345 Here we provide a complete description of an SRWM with a separate self-invited learning rate
 346 for each component. As we noted in Sec. 3, the SRWM can be split into sub-matrices: $\mathbf{W}_{t-1} =$
 347 $[\mathbf{W}_{t-1}^y, \mathbf{W}_{t-1}^q, \mathbf{W}_{t-1}^k, \mathbf{W}_{t-1}^\beta]$ according to the sub-components used to produce $\mathbf{y}_t, \mathbf{q}_t, \mathbf{k}_t$, and β_t
 348 in Eq. 5. In case where we use separate learning rates, we need separate equations to describe the
 349 update of each sub-matrices. For example, for the “y”-part \mathbf{W}_{t-1}^y , while keeping the same equation
 350 for the first projection (Eq. 5), the rest becomes:

$$\mathbf{y}_t^k = \mathbf{W}_{t-1}^y \phi(\mathbf{k}_t) \quad (9)$$

$$\mathbf{y}_t^q = \mathbf{W}_{t-1}^y \phi(\mathbf{q}_t) \quad (10)$$

$$\mathbf{W}_t^y = \mathbf{W}_{t-1}^y + \sigma(\beta_{y,t})(\mathbf{y}_t^q - \mathbf{y}_t^k) \otimes \phi(\mathbf{k}_t) \quad (11)$$

351 where \mathbf{y}_t^k and \mathbf{y}_t^q are the “y”-part of $\bar{\mathbf{v}}_t$ and \mathbf{v}_t in Eq. 6 and 7 respectively, and $\beta_{y,t} \in \mathbb{R}$ is one of four
 352 learning rate dedicated to the “y”-part. The equations for other sub-matrices $\mathbf{W}_{t-1}^q, \mathbf{W}_{t-1}^k, \mathbf{W}_{t-1}^\beta$
 353 are analogous.

354 **Use of multiple heads.** In this work, the SRWM was inserted between other layers with learned
 355 parameters with configurable dimensionalities. This allows us for an efficient computation using
 356 multiple heads for computation described as follows. Given a number of heads H which we use in
 357 the SRWM layer, we configured the model dimensions such that the input dimension to an SRWM
 358 layer d_{in} is divisible by H . The input is then split into H equally sized components, and each head
 359 executes separate SRWM operations (Eqs. 5-8) on one of the input components. In consequence, the
 360 SRWM with the same model hyper-parameters as the DeltaNet has less parameters than the DeltaNet.
 361 For example, if $d_{\text{in}} = d_{\text{key}}$, the common head dimension is $d = d_{\text{in}}/H$, the parameter shape of key
 362 projection in the SRWM is (H, d, d) while it is $(d_{\text{in}}, d_{\text{in}}) = (H * d, H * d)$ for the DeltaNet. In case
 363 the input size of the SRWM layer is not configurable, this option has to be disabled and a single head
 364 version should be used.

365 B Extra result tables

Table 3: Performance on ProcGen game environments. Multi-task training in 6 environments in the **easy distribution**. The three variants of the SRWM are as follows: *True*: the SRWM model, *Fake*: the SRWM model without self-modification mechanism, and *Reset*: the SRWM trained and evaluated with weight update reset.

Env	Split	FF	LSTM	Delta	SRM		
					True	Fake	Reset
Bigfish	Train	8.3 (3.9)	6.5 (2.0)	19.6 (4.0)	20.1 (2.4)	11.6 (5.7)	15.7 (2.8)
	Test	4.3 (2.3)	3.2 (1.1)	7.8 (1.5)	9.0 (2.0)	4.7 (2.4)	5.8 (1.3)
Fruitbot	Train	29.2 (0.2)	27.8 (0.5)	28.8 (0.9)	28.7 (0.2)	27.8 (1.3)	29.2 (0.2)
	Test	25.6 (1.1)	24.8 (0.7)	24.5 (1.5)	25.5 (1.0)	24.6 (1.2)	25.2 (1.4)
Leaper	Train	3.3 (0.2)	3.3 (0.2)	3.5 (0.4)	3.5 (0.2)	3.3 (0.3)	3.4 (0.2)
	Test	3.4 (0.4)	3.6 (0.4)	3.3 (0.4)	3.4 (0.4)	3.6 (0.4)	3.5 (0.3)
Maze	Train	1.9 (0.3)	3.1 (0.7)	3.8 (0.2)	3.6 (0.5)	3.2 (0.2)	2.9 (0.2)
	Test	1.4 (0.3)	1.6 (0.4)	1.7 (0.2)	1.8 (0.5)	1.3 (0.3)	1.5 (0.3)
Plunder	Train	3.2 (0.2)	3.2 (0.4)	3.3 (0.2)	3.1 (0.0)	3.1 (0.4)	3.1 (0.1)
	Test	3.2 (0.3)	2.9 (0.4)	3.3 (0.2)	3.0 (0.2)	3.1 (0.5)	3.0 (0.2)
Starpilot	Train	57.6 (0.9)	56.0 (1.5)	60.3 (0.4)	61.3 (2.0)	55.0 (1.3)	55.0 (1.9)
	Test	53.0 (1.7)	48.3 (2.0)	53.9 (2.4)	54.6 (2.4)	49.6 (2.1)	48.6 (1.9)
Aggregated	Train	22.5 (2.6)	28.3 (1.4)	35.0 (1.6)	27.0 (1.8)	34.6 (1.8)	28.5 (1.2)
	Test	16.4 (1.6)	15.7 (1.6)	18.6 (1.7)	20.0 (1.8)	15.3 (1.9)	16.1 (2.2)

Table 4: Performance on ProcGen game environments. Multi-task training in 4 environments in the **memory distribution**.

Env	Split	DeltaNet	SRM-Delta
Dodgeball	Train	7.1 (0.2)	7.1 (0.6)
	Test	6.4 (0.3)	6.2 (0.6)
Heist	Train	1.0 (0.3)	1.5 (0.1)
	Test	0.8 (0.2)	1.1 (0.3)
Maze	Train	5.3 (0.4)	5.9 (0.2)
	Test	3.3 (0.6)	3.3 (0.4)
Miner	Train	32.3 (0.4)	34.5 (0.8)
	Test	29.2 (1.1)	29.4 (0.7)
Aggregated	Train	51.8 (2.6)	59.0(2.1)
	Test	38.0 (4.1)	38.5(3.2)