

## Supplementary Material

### A ADDITIONAL RESULTS

Figures A1 – A12 show heatmaps of mean best action set density for XCS and PPL-ST across all FL environments after 250 epochs. This is an extension of the results presented in Figure 3 of the main paper, although note that Figure 3 is equivalent to Figure A7a with the addition of relevant information from Table A1. Each of these figures (excluding A12) shows two heatmaps, representing the same data on two different scales: local, which is environment-specific (left-hand subfigure) and global, which is shared across all environments (right-hand subfigure). The minimum value on all scales was set to 1. The maximum value on the global scale was set based on the maximum value represented on any of the heatmaps; this occurring in the (12, 0.5) environment. Therefore there is only a single heatmap for the (12, 0.5) environment in Figure A12, as the local and global scales for this environment are identical. The intention of using both scales is to clearly visualise the patterns present in specific environments as well as across all environments.

**Table A1: Ruleset size statistics for XCS and PPL-ST across all FL environments after 250 epochs.** For both systems,  $|R|$  represents the size of the entire ruleset evolved by the system, and  $|R_{BA}|$  represents the size of the set of rules comprising the best action map. For XCS both of these quantities are measured in units of number of macroclassifiers. Overbars indicate sample means, computed over thirty trials. Note that  $|R|$  for PPL-ST is not a sample mean because by definition this quantity is equivalent to the *idvSize* hyperparameter, which is constant for each grid size (see experimental setup in Section 6 of main paper).

		XCS		PPL-ST	
		$\overline{ R }$	$\overline{ R_{BA} }$	$\overline{ R }$	$\overline{ R_{BA} }$
FL env.	(4, 0)	144.0	63.8	7	6.0
	(4, 0.1)	179.4	68.9		6.7
	(4, 0.3)	194.9	83.4		6.5
	(4, 0.5)	202.5	108.2		6.5
	(8, 0)	521.4	261.9	21	17.8
	(8, 0.1)	594.6	236.0		17.4
	(8, 0.3)	615.7	272.2		18.4
	(8, 0.5)	591.7	321.0		18.4
	(12, 0)	1124.0	539.8	42	33.1
	(12, 0.1)	1264.2	527.3		34.5
	(12, 0.3)	1458.9	697.6		35.2
	(12, 0.5)	1303.1	790.4		34.1

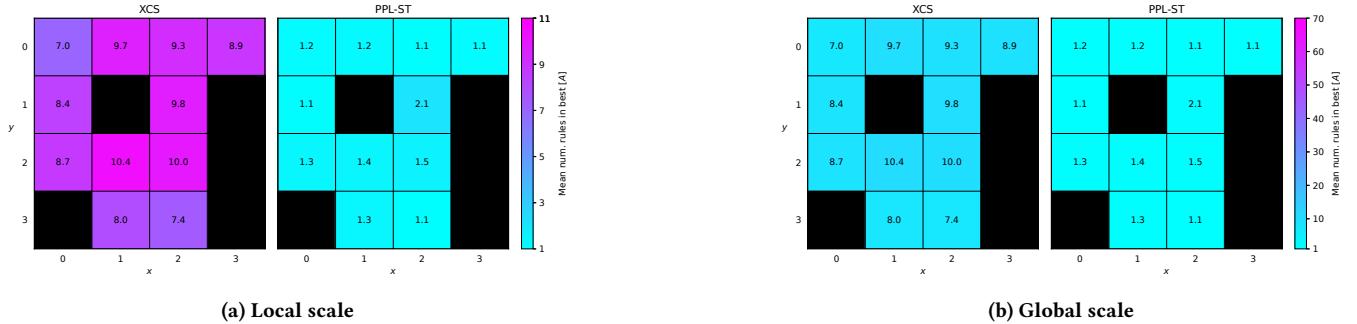


Figure A1: Mean best action set density of XCS and PPL-ST in (4, 0) FL after 250 epochs.

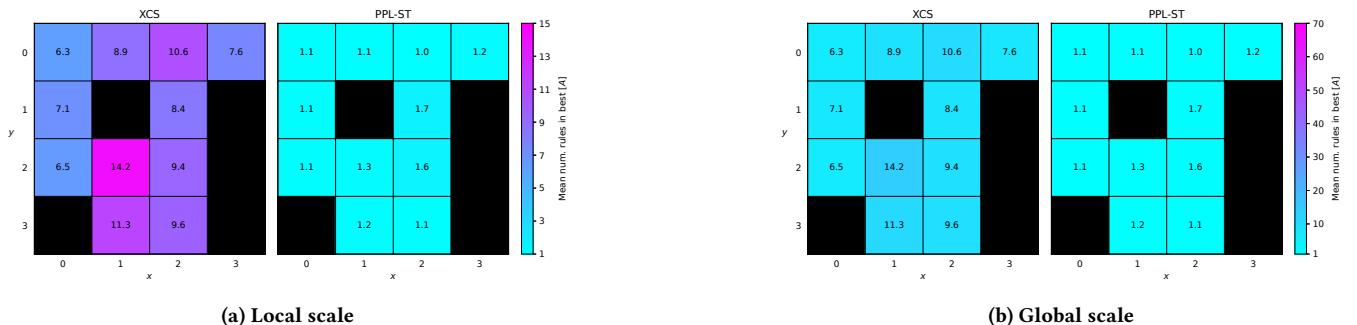


Figure A2: Mean best action set density of XCS and PPL-ST in (4, 0.1) FL after 250 epochs.

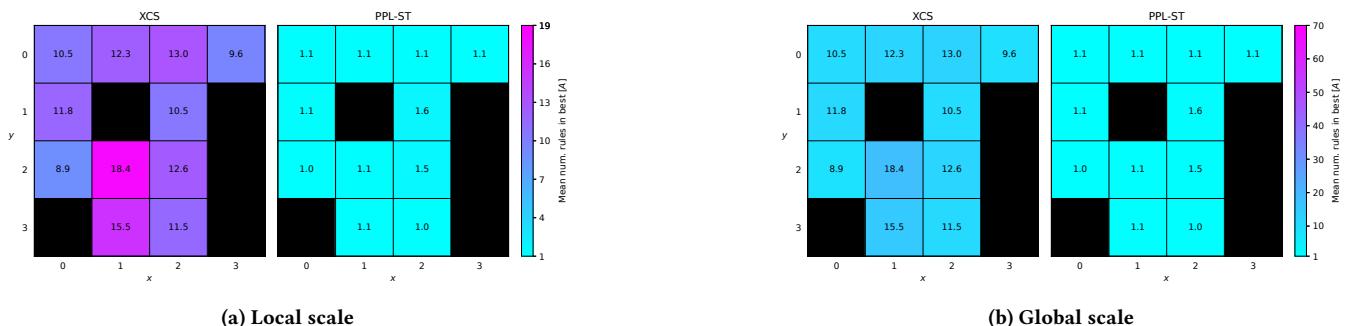


Figure A3: Mean best action set density of XCS and PPL-ST in (4, 0.3) FL after 250 epochs.

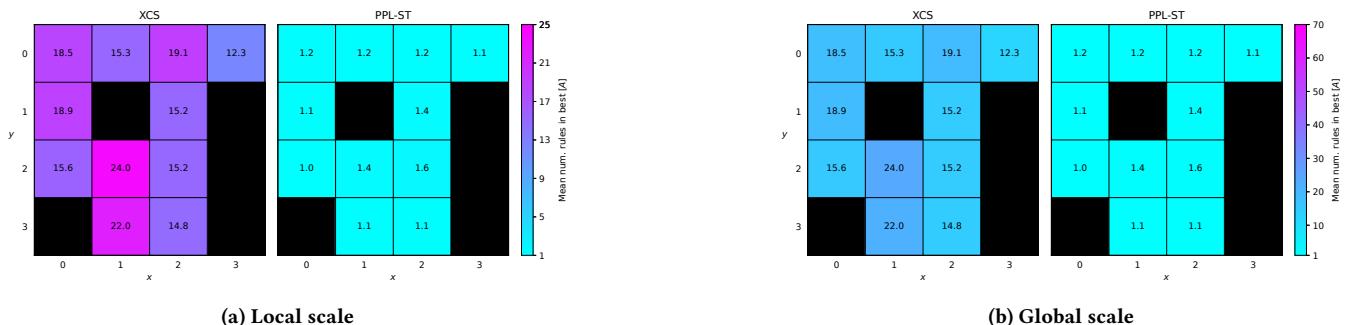


Figure A4: Mean best action set density of XCS and PPL-ST in (4, 0.5) FL after 250 epochs.

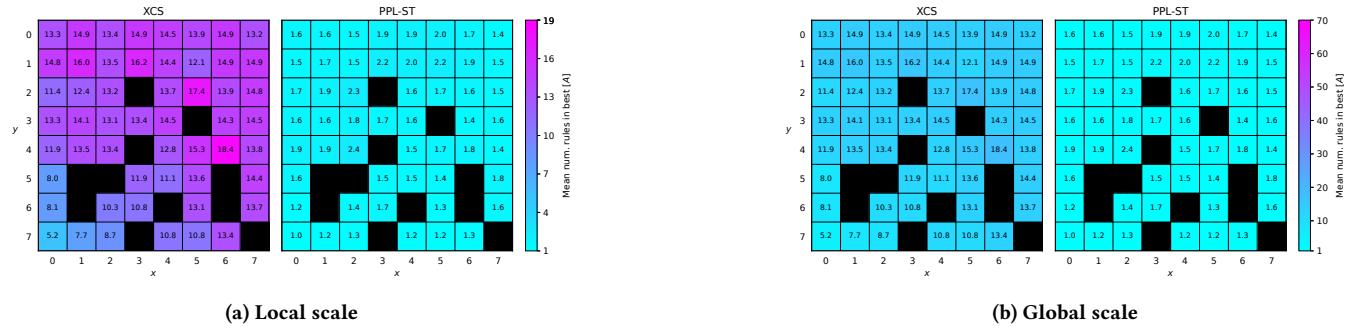


Figure A5: Mean best action set density of XCS and PPL-ST in (8, 0) FL after 250 epochs.

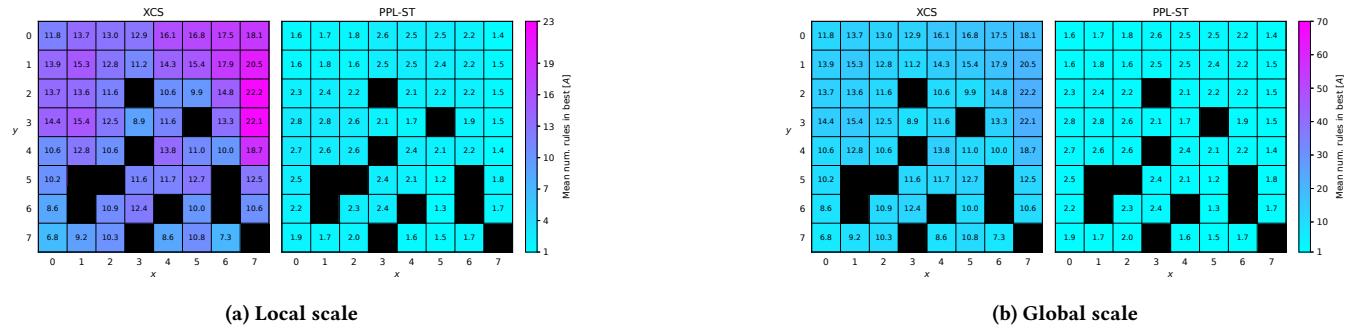


Figure A6: Mean best action set density of XCS and PPL-ST in (8, 0.1) FL after 250 epochs.

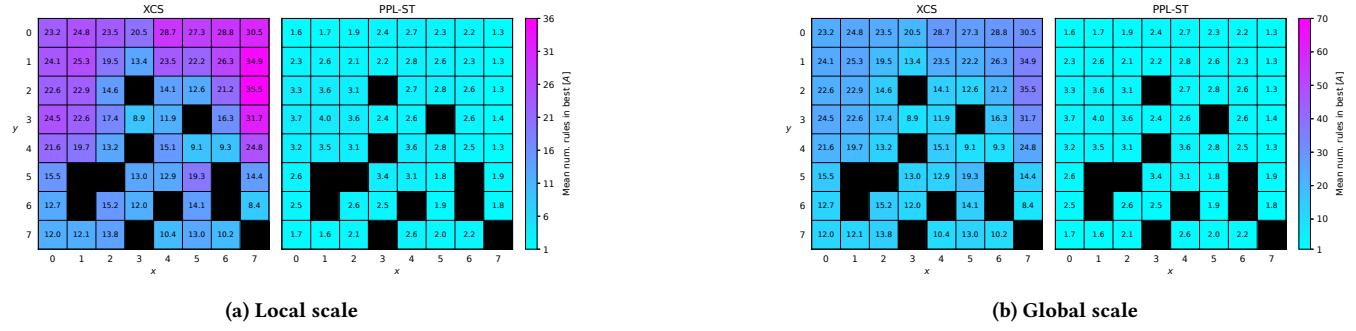


Figure A7: Mean best action set density of XCS and PPL-ST in (8, 0.3) FL after 250 epochs.

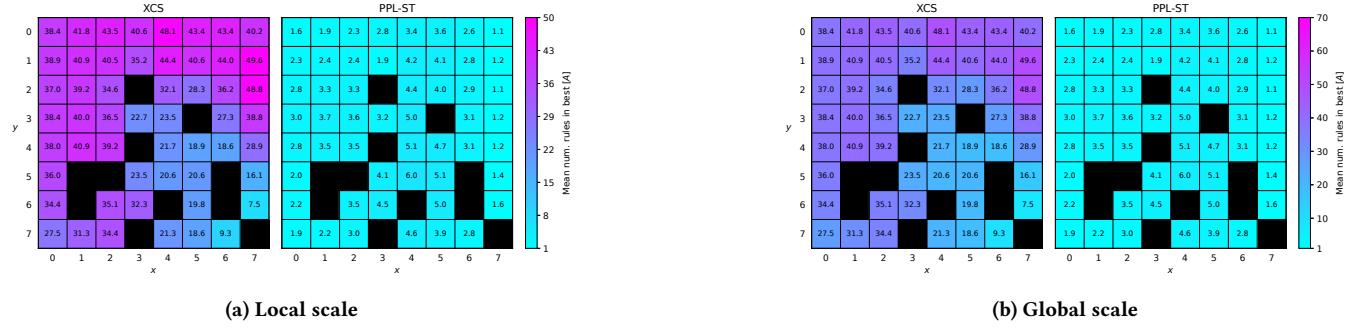


Figure A8: Mean best action set density of XCS and PPL-ST in (8, 0.5) FL after 250 epochs.

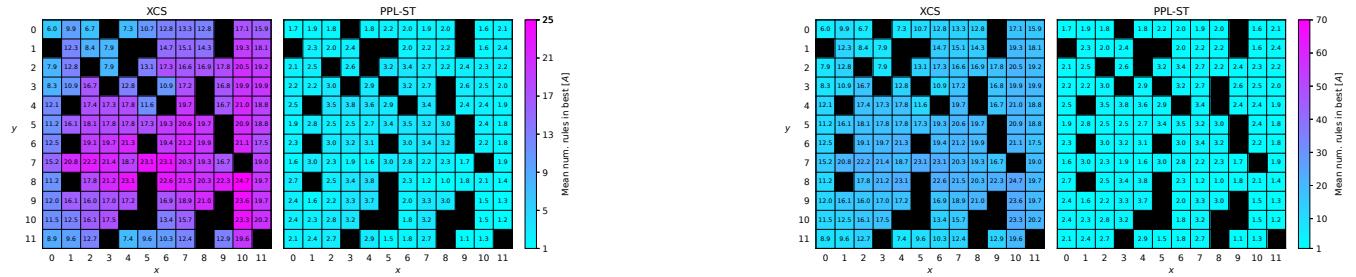


Figure A9: Mean best action set density of XCS and PPL-ST in (12, 0) FL after 250 epochs.

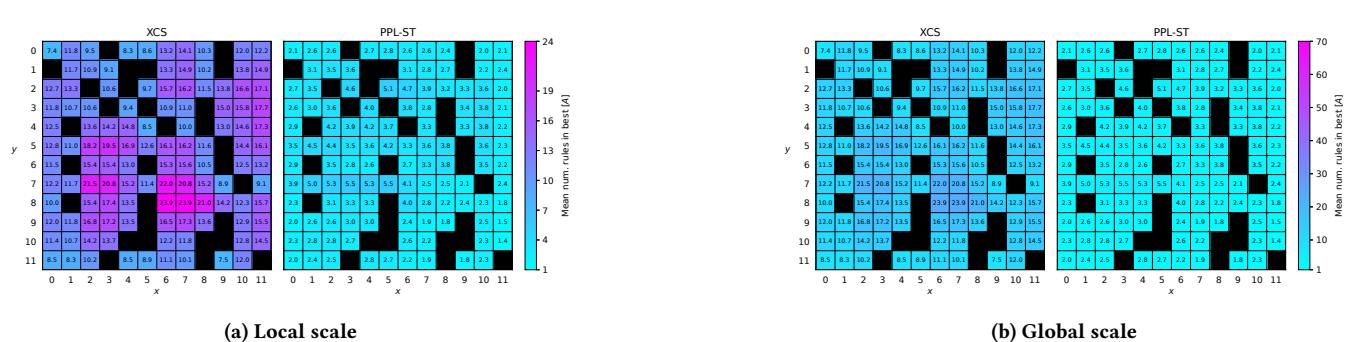


Figure A10: Mean best action set density of XCS and PPL-ST in (12, 0.1) FL after 250 epochs.

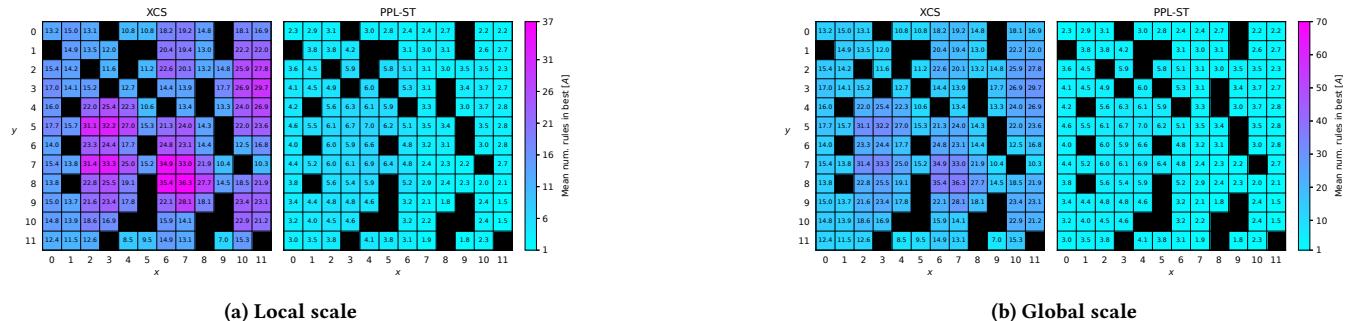


Figure A11: Mean best action set density of XCS and PPL-ST in (12, 0.3) FL after 250 epochs.

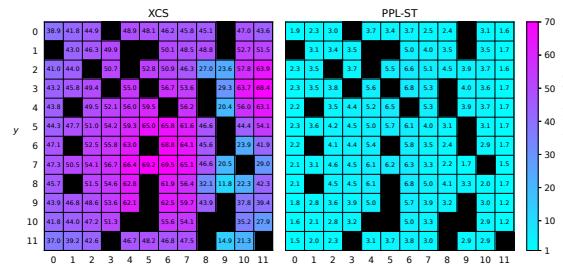


Figure A12: Mean best action set density of XCS and PPL-ST in (12, 0.5) FL after 250 epochs.