PSRL samples PriorMDP ($t = 500$, $N_s = 4$, $N_a = 2$)

Legend: ✗ Sample mean (red) ✗ True $Q_{\mathbf{s},\mathbf{a}}^*$ (black)