Figure 3: Sensitivity of sparse mirror-descent TD to noisy features in a grid-world domain. Left: basis matrix with the first 50 columns representing proto-value function bases and the remainder 450 bases representing mean-0 Gaussian noise. Right: Approximated value function using sparse mirror-descent TD.

Figure 4: Left: convergence of mirror-descent Q-learning with a fixed p-norm link function. Right: decaying p-norm link function.

Figure 5: Left: Convergence of composite mirror-descent Q-learning on two-room gridworld domain. Right: Approximated value function, using 50 proto-value function bases.

Figure 6: Top: Q-learning; Bottom: mirror-descent Q-learning with p-norm link function, both with 25 fixed Fourier bases [KOT08] for the mountain car task.

6 Comparison of Link Functions

The two most widely used link functions in mirror descent are the p-norm link function [BT03] and the relative entropy function for exponentiated gradient (EG) [KW95]. Both of these link functions offer a multiplicative update rule compared with regular additive gradient methods. The differences between these two are discussed here. Firstly, the loss function for EG is the relative entropy whereas that of the p-norm link function is the square $l_2$-norm function. Second and more importantly, EG does not produce sparse solutions since it must maintain the weights away from zero, or else its potential (the relative entropy) becomes unbounded at the boundary.

Another advantage of $p$-norm link functions over EG is that...