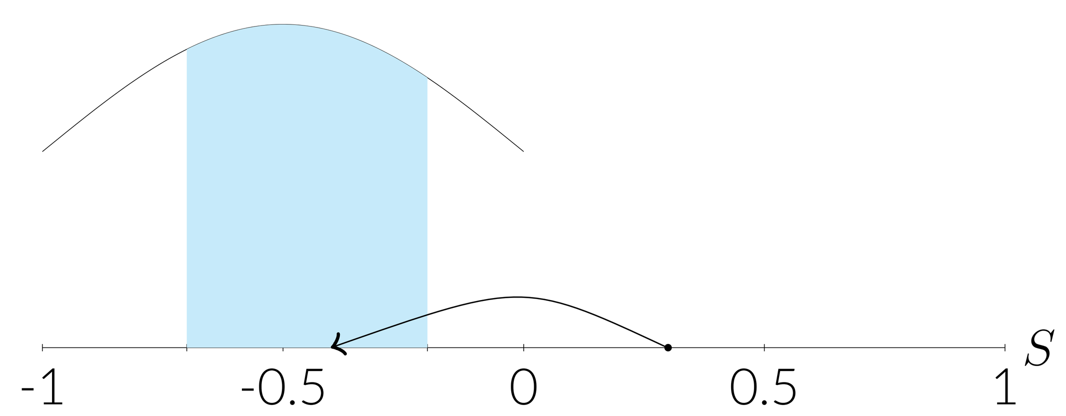
Reachability in Uncountable MDPs

K. Grover J. Křetínský T. Meggendorfer M. Weininger

TU Munich

Uncountable MDP



$\mathcal{M} = (S, Act, Av, \Delta)$

Reachability Problem

Find the probability of reaching the target area from a given starting point in an MDP [2] with continuous state space.

We give an anytime algorithm which approximates the value function with a converging bound on the error.

Assumptions

- Lipschitz Continuity
- State-Action Maximum Appoximation
- Transition Approximation
- State-Action Sampling
- Sink Computability and Attractor

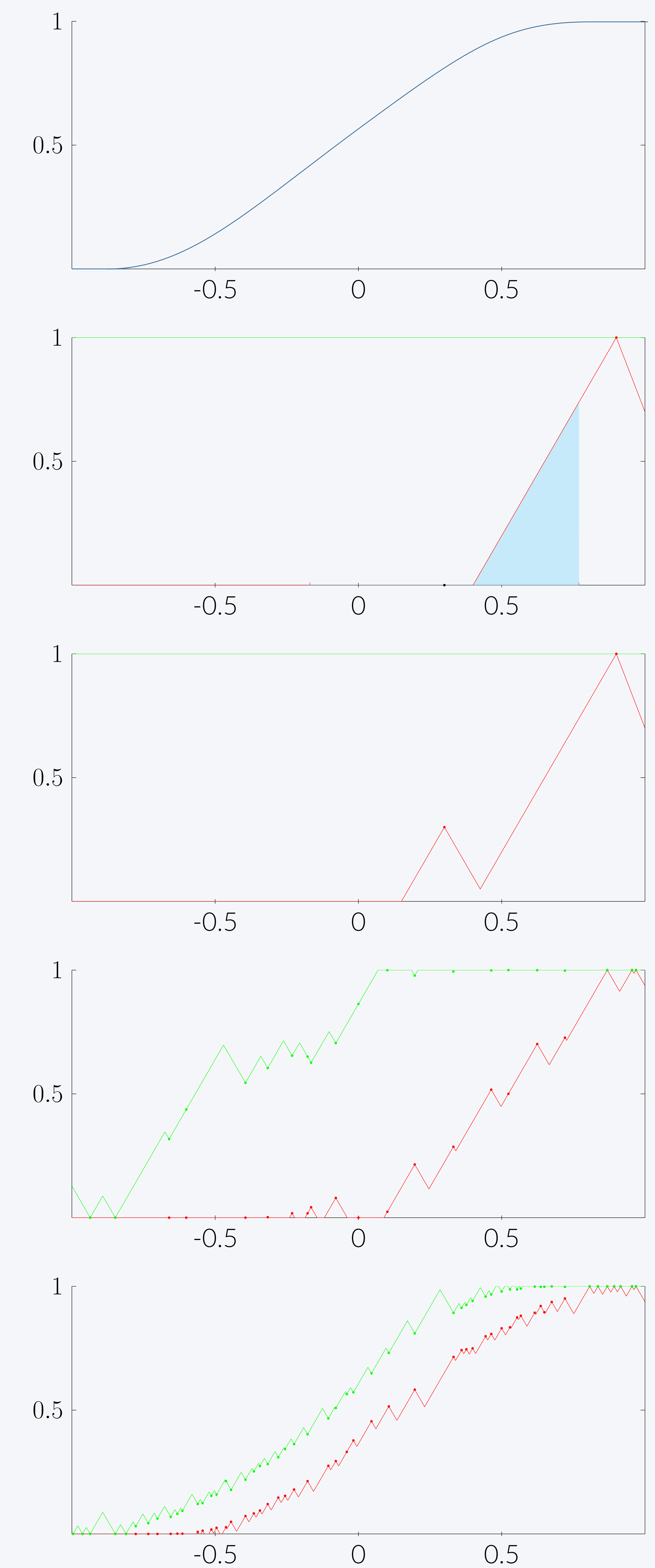
Overview

Bounded Real Time Dynamic Programming (BRTDP) [1] solves reachability problem in the finite setting very efficiently. Our idea is to extend BRTDP to the uncountable setting.

The following are intuitions behind the assumptions:

- If we can sample fairly from the whole state space i.e. always eventually sample from an ϵ neighborhood of all the points for any $\epsilon > 0$ only then we would be able to approximate the value function arbitrarily close.
- Using Lipschitz continuity, if we know a nontrivial lower bound at some point of the state space then we can safely approximate the lower bounds on the neighbouring points as well as shown in Figure 1.

Algorithm



Overview

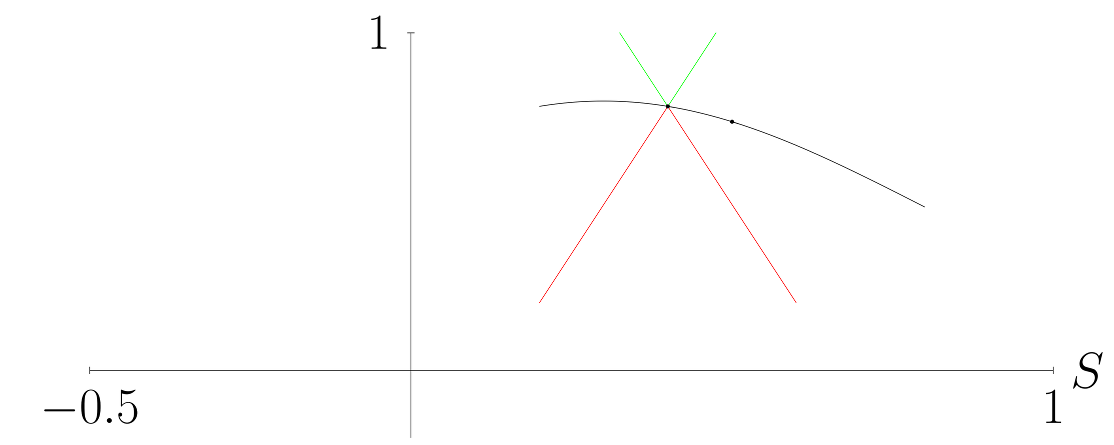
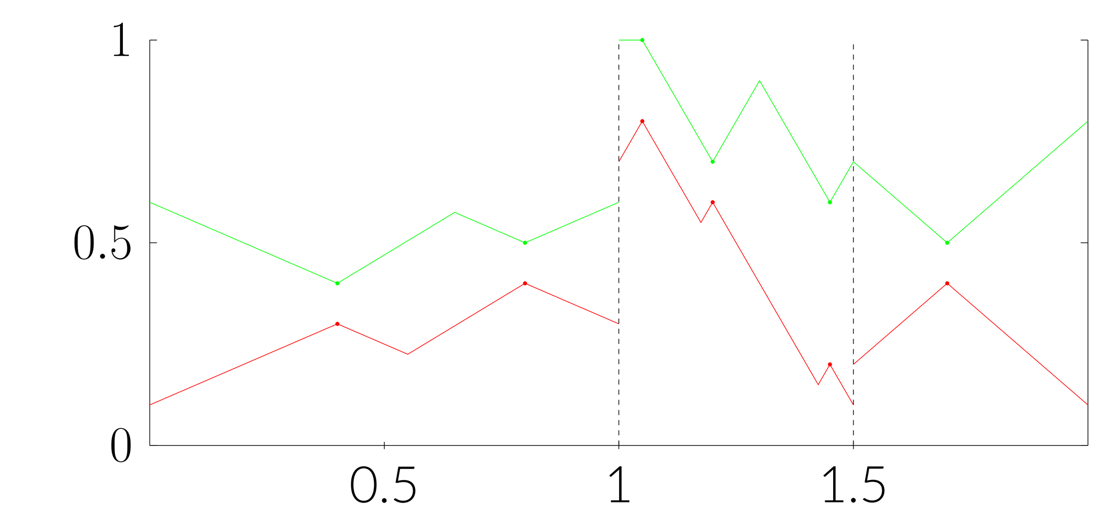


Figure 1:Lipschitz continuity

- With a target set T, we need another sink R from which the probability of reaching T is 0. T and R helps in updating lower and upper bounds respectively.

Extensions

- Discontinuities:



- **LTL:** Can handle "reach-avoid" properties directly.
- **Apply learning:** Just like the BRTDP in finite case, we can use learning heuristics to guide the algorithm here as well. For e.g. a Learning procedure would predict which area of the state space is important and then the algorithm would try to sample points from that area with large probability.

References

[1] Tomáš Brázdil, Krishnendu Chatterjee, Martin Chmelík, Vojtěch Forejt, Jan Křetínský, Marta Kwiatkowska, David Parker, and Mateusz Ujma.
Verification of markov decision processes using learning algorithms.
In Franck Cassez and Jean-François Raskin, editors, *Automated Technology for Verification and Analysis*, pages 98--114, Cham, 2014. Springer International Publishing.

[2] Martin L. Puterman.
Markov Decision Processes: Discrete Stochastic Dynamic Programming.
Wiley Series in Probability and Statistics. Wiley, 1994.