



THE UNIVERSITY OF
MELBOURNE

Scalable Solutions to Zero-Sum Partially Observable Stochastic Games Through Belief Aggregation with Approximation Guarantees

40th AAAI Conference on Artificial Intelligence

Singapore, January 20-27 2026

Kim Hammar and Tansu Alpcan

{kim.hammar, tansu.alpcan}@unimelb.edu.au

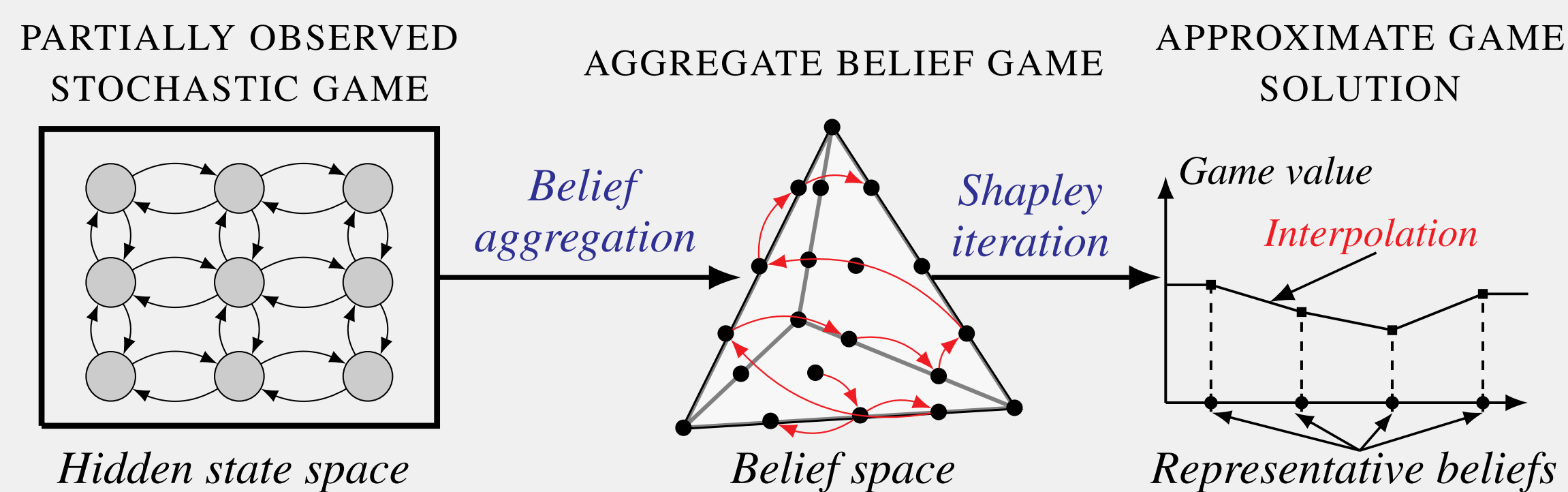
The University of Melbourne, Department of Electrical and Electronic Engineering



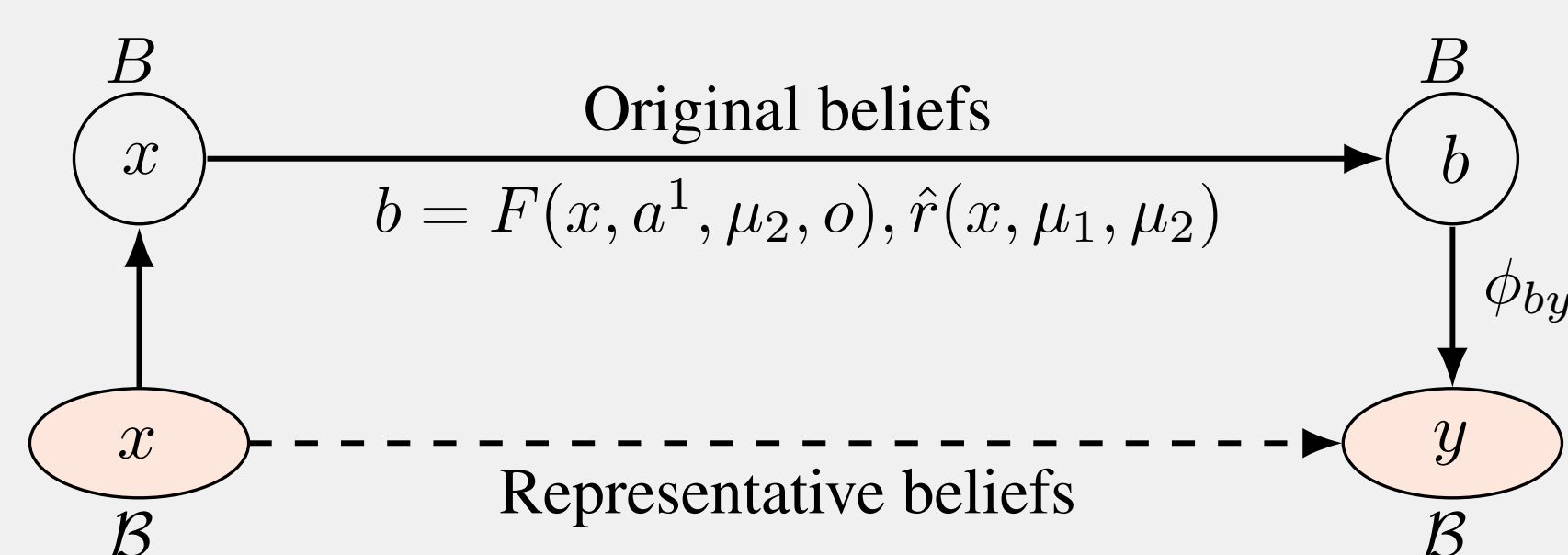
Contributions

1. We present SAB, a **scalable method for solving one-sided zero-sum POSGs**.
2. We establish a bound on the approximation error and prove convergence.
3. We show that SAB **outperforms the state-of-the-art on several games**.

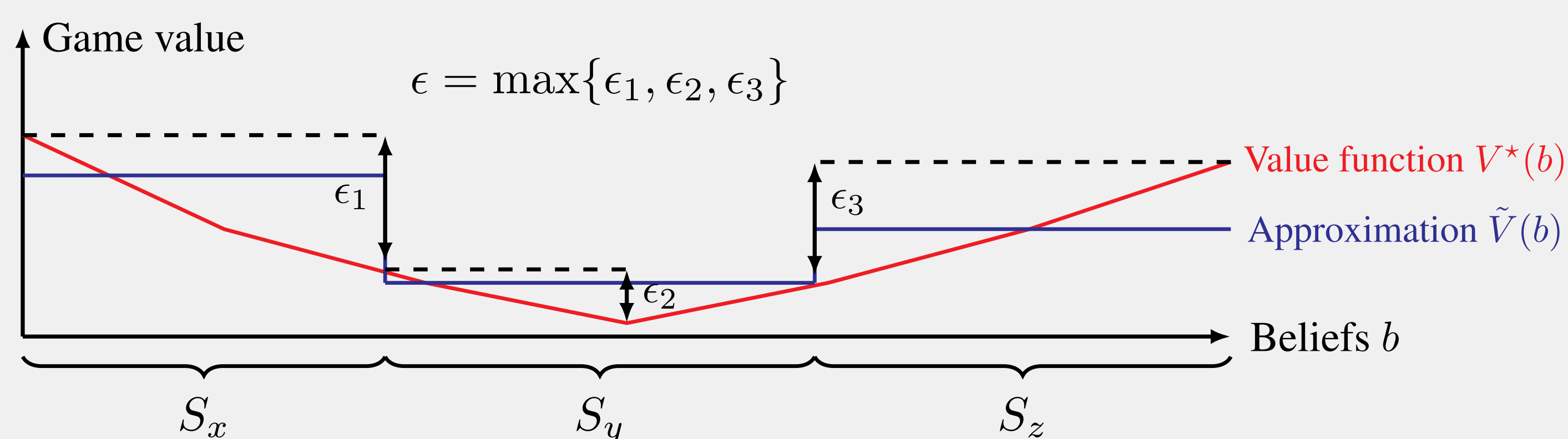
Shapley iteration with Aggregated Beliefs (SAB)



Aggregate Belief Game



Value Function Approximation



Main Theoretical Results (Informal)

Proposition 1 (Convergence). SAB converges for any one-sided zero-sum POSG. (Contraction argument based on Banach fixed-point theorem.)

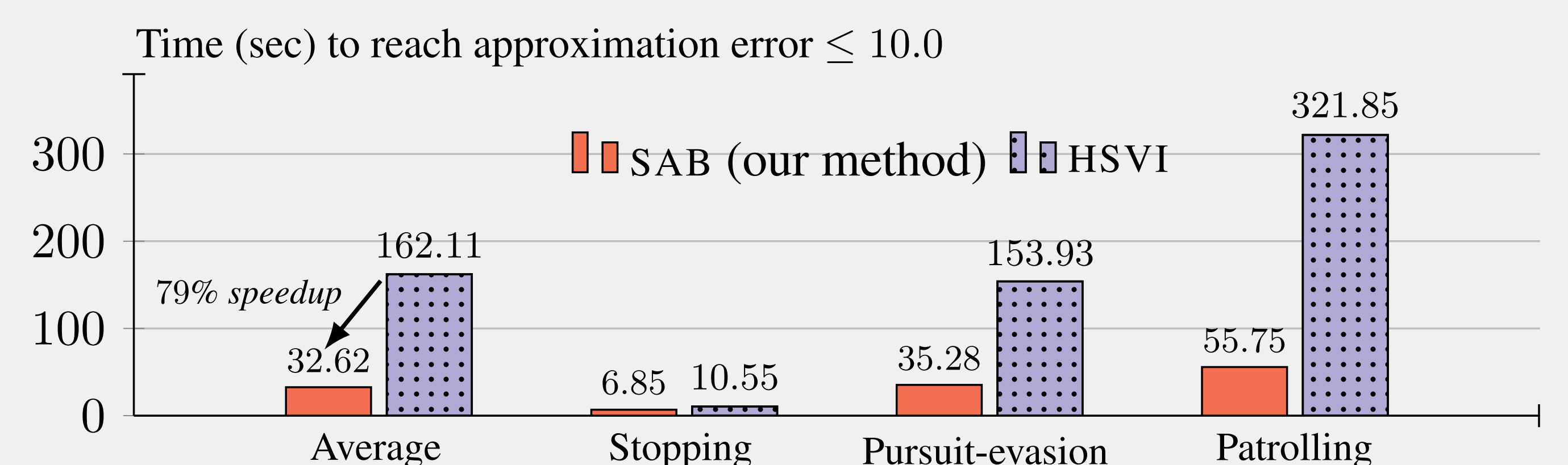
Proposition 2 (Approximation error bound). The difference between the estimated value function \tilde{V} and the true value function V^* is bounded as

$$|\tilde{V}(b) - V^*(b)| \leq \frac{\epsilon}{1 - \gamma}, \quad \text{for all } b \in B,$$

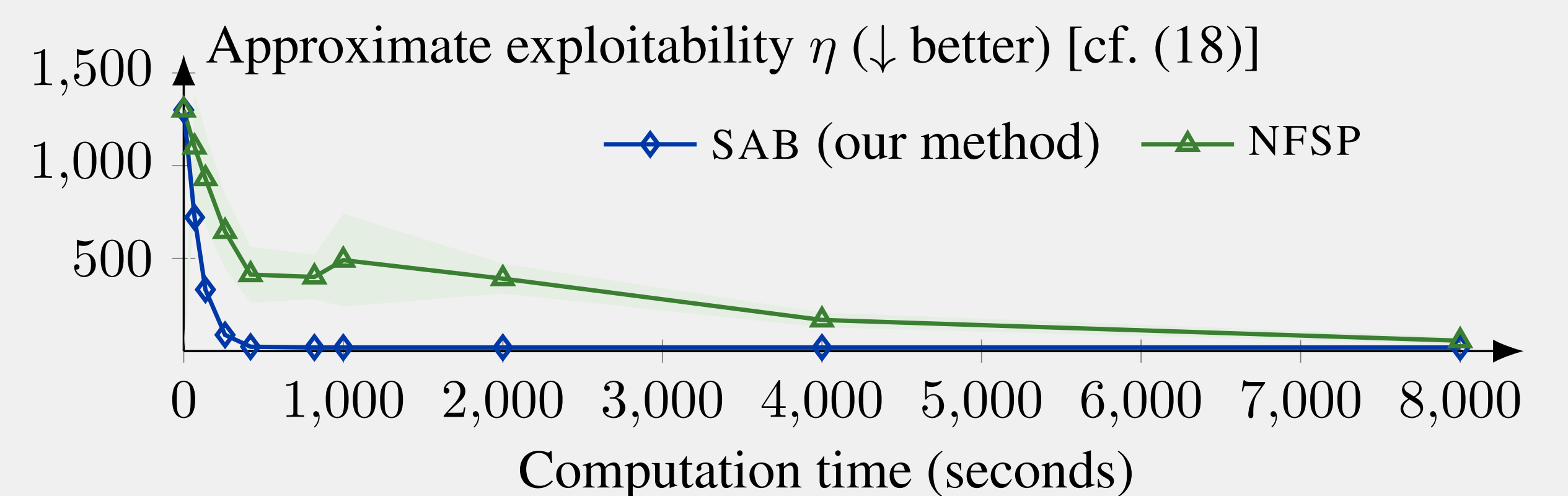
where γ is the discount factor and ϵ is a finite constant defined by

$$\epsilon = \max_{x \in \mathcal{B}} \sup_{b, b' \in S_x} |V^*(b) - V^*(b')|, \quad S_x = \{b \mid b \in B, \phi_{bx} = 1\}.$$

Scalability



Comparison with Neural Fictitious Self-Play (NFSP)



Numerical illustrations of Theoretical Results

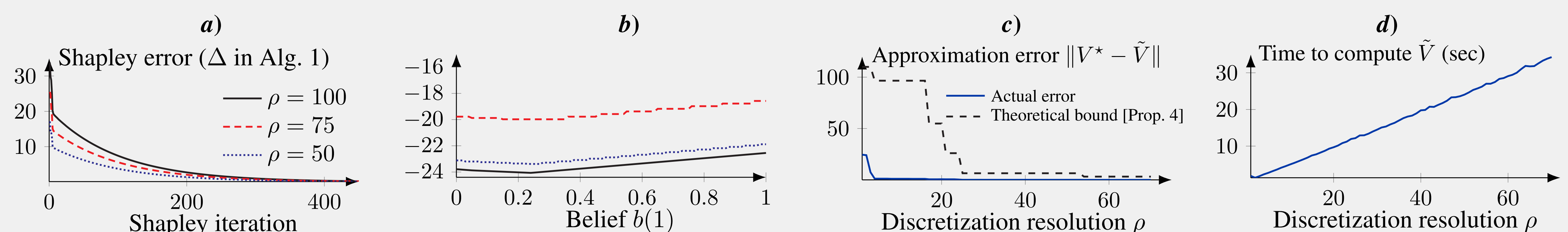


Figure: Numerical illustrations on a small game with two states. Plot **a)** shows the convergence of SAB for varying discretization resolutions of the belief space; plot **b)** visually compares the approximation obtained through SAB with the value function; plot **c)** compares the approximation error with the theoretical bound; and plot **d)** shows the compute time.

Comparison with Heuristic Search Value Iteration (HSVI)

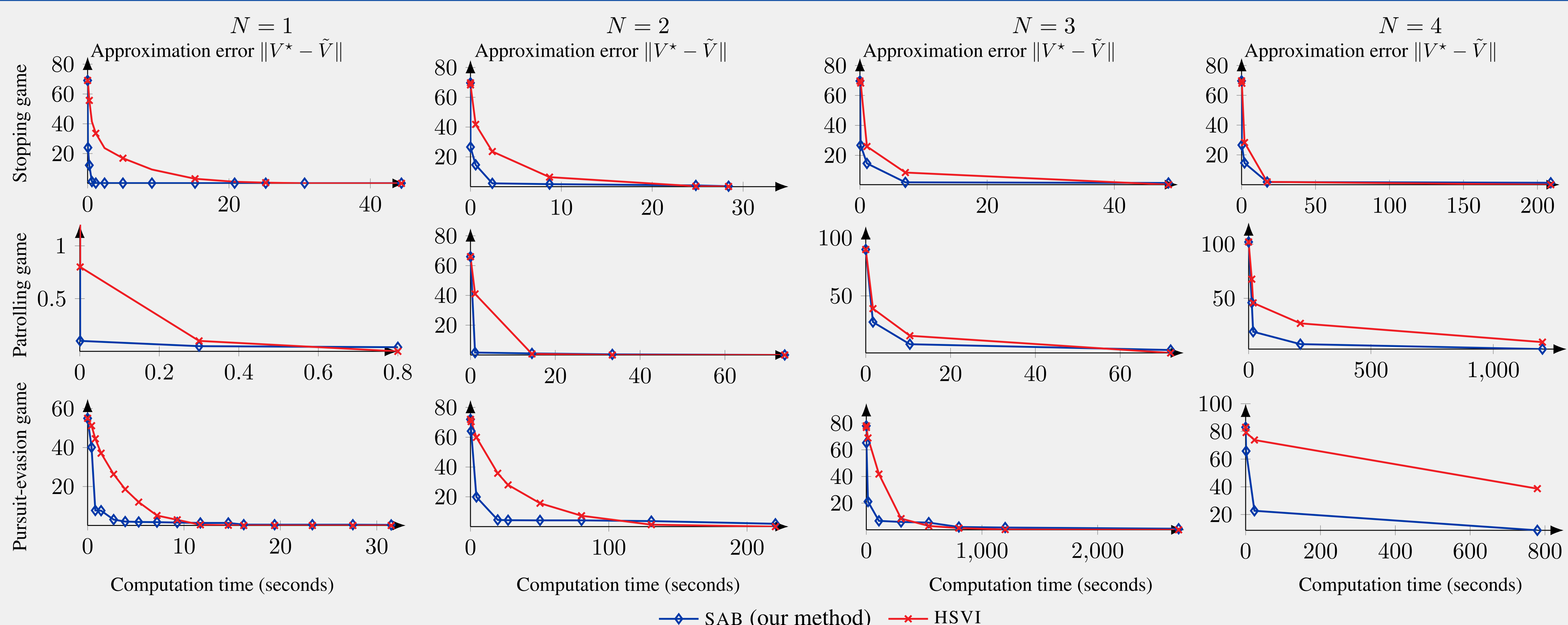


Figure: Comparison between our method (SAB) and the current state-of-the-art method (HSVI) on three example POSGs from the game-theoretic literature. Rows relate to different games. Columns relate to the parameter N , which controls the size of the game instantiation. The x-axes indicate computation time and the y-axes indicate the approximation error.