# Adapting Beyond the Depth Limit: Counter Strategies in Large Imperfect Information Games

Extended Abstract

David Milec
AI Center, FEE, CTU in Prague
Czech Republic
milecdav@fel.cvut.cz

Vojtěch Kovařík
AI Center, FEE, CTU in Prague
Czech Republic
vojta.kovarik@gmail.com

Viliam Lisý
AI Center, FEE, CTU in Prague
Czech Republic
viliam.lisy@agents.fel.cvut.cz

## ABSTRACT

We investigate how to adapt against known sub-optimal opponents while maintaining robustness against rational players in large imperfect-information zero-sum games. Previous approaches to large games use depth-limited search since examining the complete game tree is computationally infeasible. In computing a robust strategy against an opponent, the latest methods assume rational play beyond the search depth limit, restricting their ability to adapt to the opponent's behavior. To address this limitation, we introduce *Adapting Beyond Depth-limit* (ABD). This algorithm employs a strategy-portfolio approach – which is called matrix-valued states – for depth-limited search. ABD is the first robust adaptation method capable of fully utilizing all available information about opponent models in large imperfect-information games. The matrix-valued states approach also simplifies the algorithm compared to previous methods that rely on optimal value functions. Our experiments demonstrate that ABD may double the utility when facing opponents who make mistakes beyond the depth and significantly improves utility against randomly generated opponents while maintaining safety against worst-case rational adversaries.

## KEYWORDS

opponent adaptation; imperfect information; best response; depth-limited search

**ACM Reference Format:**
David Milec, Vojtěch Kovařík, and Viliam Lisý. 2025. Adapting Beyond the Depth Limit: Counter Strategies in Large Imperfect Information Games: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

## 1 INTRODUCTION

Recent advances in solving large imperfect-information games rely on depth-limited methods that avoid searching through intractably large full game trees [1–3, 12]. Instead, these methods look a few steps ahead and use value functions, typically represented by neural networks, to approximate outcomes beyond the depth limit [7].

Most existing work assumes fully rational opponents, but this assumption often fails in practice due to cognitive or computational limitations [4, 9, 13]. While some research addresses sub-rational opponents in depth-limited settings, current methods cannot exploit opponent's mistakes beyond the depth limit [5, 8, 11]. For instance, if an opponent has an exploitable behavior that occurs $d + 1$ steps ahead, a method that only looks $d$ steps ahead while assuming optimal play thereafter will miss this opportunity.

We introduce Adapting Beyond Depth-limit (ABD), which uses matrix-valued states instead of optimal value functions at the depth limit. This approach allows players to choose from a strategy portfolio, with utilities determined by their joint choices. ABD adapts to sub-rational opponents by replacing their portfolio with their modeled strategy, making it the first depth-limited method capable of utilizing all opponent mistakes while maintaining robustness.

Experimental results in poker and battleship show that ABD yields more than a twofold increase in utility when facing opponents who make mistakes beyond the depth limit and also delivers significant improvements in utility against randomly generated opponents.

The full version of the paper can be found at [10].



*Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA.* © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

## 2 BACKGROUND

Our method is based on the Restricted Nash response [6], which creates robust counter-strategies in two-player games. It assumes there's a probability $p$ that the opponent will use their fixed strategy and a probability *(1-p)* that they'll play rationally knowing the strategy we deployed. The parameter $p$ effectively controls how much the strategy attempts to adapt to the fixed opponent strategy versus playing safely against the worst-case adversary.

This scenario can be modeled by creating a modified version of the original game, starting with a chance node. Based on this chance event, which only the opponent can observe, the game proceeds in one of two ways: either the opponent is forced to play their fixed strategy, or both players play the original game without restrictions. The optimal strategy for this modified game is a **$p$-restricted Nash response** to the opponent's fixed strategy [6].

## 3 ADAPTATION BEYOND THE DEPTH-LIMIT

In this section, we describe the method *adapting beyond depth-limit* (ABD) that can strike a balance between playing well against a specific opponent and remaining unexploitable. Unlike previous depth-limited methods, ABD can take advantage of the opponent's mistakes, irrespective of whether they occur within or beyond the depth limit. In Section 3.1, we describe the idealised version of the method, assuming we have a value function that captures the behavior of the specific opponent. In Section 3.2, we describe two practical methods for approximating the idealised value function described in Section 3.1.

### 3.1 Idealised ABD

The idealised version of the Adapting Beyond Depth-limit (ABD) algorithm operates on games where each player has a set of possible strategies (called a portfolio). The algorithm takes three parameters: a depth limit determining how far ahead to look in the game tree, a fixed strategy representing the opponent's expected behavior, and a probability parameter that balances adaptation to the opponent and maintaining robustness. The algorithm works by transforming the original game through the following steps:

First, it creates a modified game that implements the restricted Nash response mechanism, allowing for either facing the fixed opponent strategy or a rational opponent. Then, it creates a subgame starting from the current game state, incorporating the information about previously played moves, and uses a gadget to ensure robustness [11]. Finally, it creates a depth-limited version of this subgame where players must choose from their strategy portfolios after reaching the depth limit. In the idealised version, we assume the portfolio consists of all the pure strategy continuations.

The algorithm then finds a Nash equilibrium in this transformed game and uses it to determine the next action. This algorithm is idealised since when constructing the depth-limited version, we need the values of the portfolio against the fixed opponent strategy. In the idealised version, we assume we have access to those. However, in practical implementation, we need to compute them, and we discuss that in Section 3.2

The gadget at the top of the modified game follows the construction described by [11]. In the subgame where the opponent is fixed, we introduce an initial chance node, which sets the initial reaches at the root of the subgame to the combined reaches of all players, including chance. In the subgame where the opponent plays rationally, we use the full gadget proposed by [11]. In this idealised case, we show that ABD produces a $p$-restricted Nash response [10].

### 3.2 Practical ABD Implementation

The idealised version faces two practical limitations in large games. First, it's impractical to include all pure strategies in the portfolios. For example, even a modest-sized Battleships game with a 5x5 board and two 2x1 ships would require portfolios containing approximately $10^{25}$ strategies. This is addressed by using a limited selection of diverse strategies.

Second, computing exact values for all strategy combinations is computationally expensive, even with moderately sized portfolios. While we can precompute values when both players use portfolio strategies, we face an opponent we need to model in practice. Therefore, the fixed opponent remains unknown until gameplay begins. Therefore, values against this opponent must be estimated during play. We use a sampling approach: for each leaf history, we sample $n$ continuation trajectories for every strategy in our portfolio, assuming our agent follows the chosen strategy from the portfolio and the opponent follows its fixed strategy. Then, we use the average of these sampled values as our estimate.

These practical considerations lead to an implementation that balances theoretical correctness with computational feasibility while maintaining the algorithm's core ability to adapt to opponent behavior beyond the depth limit.

## 4 EXPERIMENTS

We conduct experiments in both small and large imperfect information games to highlight the failure modes of continual depth-limited best/robust response (CDBR/CDRNR) [11] and demonstrate the adaptability of our method against subrational strategies.

### 4.1 Failure of Previous Methods

In our first experiment, using a 2x2 Battleships game with a 1x1 ship, we tested against an opponent who shoots uniformly except for targeting the top-left corner as their last move. CDBR failed to exploit this pattern at all, only achieving the game's base value (0.25) until observing two out of three significant moves. In contrast, ABD consistently won using a portfolio of four strategies, each avoiding a different cell while shooting uniformly elsewhere.

| **Depth** | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **CDBR** | 0.25 | 0.25 | 1.00 | 1.00 |
| **ABD** ($p = 1$) | 1.00 | 1.00 | 1.00 | 1.00 |

**Table 1: Winrate at different depths on 2x2 battleships with one 1x1 ship against an opponent who always shoots the top left corner last. Depth is the number of future opponents' moves contained in the depth-limited subgame.**

### 4.2 Leduc Experiments

We then conducted experiments in Leduc poker against four specific opponent strategies and 1,000 randomly generated ones. The opponent strategies (S1-S4) included variations of aggressive/passive patterns in different rounds. ABD significantly outperformed CDBR against both the specific strategies and random opponents, though the improvement margin was smaller against random strategies (with 95% confidence intervals).

We compared ABD with CDRNR against S1. CDRNR performs better for very small $p$ because it uses the optimal value function. For $p > 0.2$ ABD beats CDRNR and is very close to the RNR.

| | S1 | S2 | S3 | S4 | Random |
|---|---|---|---|---|---|
| **CDBR** | 1 | 5 | 1 | 3 | 2.475 ± 0.016 |
| **ABD** ($p = 1$) | 2.3 | 5 | 4.2 | 5 | 2.536 ± 0.015 |

**Table 2: Results of CDBR and ABD against heuristic strategies S1 to S4 and Random in Leduc hold'em.**

## 5 CONCLUSION

This paper tackled adapting to subrational opponents in large, imperfect-information games with depth-limited solving. We introduced a novel framework using matrix-valued states to model opponent strategies beyond the depth limit. Our method demonstrated slightly better robust adaptation to random opponents compared to CDRNR, the state-of-the-art approach. Notably, it significantly outperformed CDRNR against opponents making mistakes in later game stages, achieving a twofold increase in **gain**.

## ACKNOWLEDGMENTS

This research was supported by the Czech Science Foundation (grant no. GA22-26655S and grant no. GA25-18353S).

## REFERENCES


[1] Noam Brown and Tuomas Sandholm. 2018. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359, 6374 (2018), 418–424.

[2] Noam Brown and Tuomas Sandholm. 2019. Superhuman AI for multiplayer poker. *Science* 365, 6456 (2019), 885–890.

[3] Neil Burch, Michael Johanson, and Michael Bowling. 2014. Solving imperfect information games using decomposition. In *Twenty-eighth AAAI conference on artificial intelligence*. AAAI Press, Quebec, 602–608.

[4] Fei Fang, Thanh Hong Nguyen, Rob Pickles, Wai Y Lam, Gopalasamy R Clements, Bo An, Amandeep Singh, Brian C Schwedock, Milind Tambe, and Andrew Lemieux. 2017. PAWS-A Deployed Game-Theoretic Application to Combat Poaching. *AI Magazine* 38, 1 (2017), 23–36.

[5] Zhenxing Ge, Zheng Xu, Tianyu Ding, Linjian Meng, Bo An, Wenbin Li, and Yang Gao. 2024. Safe and Robust Subgame Exploitation in Imperfect Information Games. In *Forty-first International Conference on Machine Learning*. Proceedings of Machine Learning Research, Vienna, 15255 – 15270.

[6] Michael Johanson, Martin Zinkevich, and Michael Bowling. 2008. Computing robust counter-strategies. In *Advances in neural information processing systems*. 721–728.

[7] Vojtěch Kovařík, Dominik Seitz, Viliam Lisý, Jan Rudolf, Shuo Sun, and Karel Ha. 2023. Value functions for depth-limited solving in zero-sum imperfect-information games. *Artificial Intelligence* 314 (2023), 103805.

[8] Mingyang Liu, Chengjie Wu, Qihan Liu, Yansen Jing, Jun Yang, Pingzhong Tang, and Chongjie Zhang. 2022. Safe Opponent-Exploitation Subgame Refinement. In *Advances in Neural Information Processing Systems*. 27610 – 27622.

[9] David Milec, Jakub Černý, Viliam Lisý, and Bo An. 2021. Complexity and Algorithms for Exploiting Quantal Opponents in Large Two-Player Games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 5575–5583.

[10] David Milec, Vojtěch Kovařík, and Viliam Lisý. 2025. Adapting Beyond the Depth Limit: Counter Strategies in Large Imperfect Information Games. *arXiv preprint arXiv:2501.10464* (2025).

[11] David Milec, Ondřej Kubíček, and Viliam Lisý. 2024. Continual Depth-limited Responses for Computing Counter-strategies in Sequential Games. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. 2393–2395.

[12] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. 2017. Deepstack: Expert-level artificial intelligence in Heads-Up No-Limit Poker. *Science* 356, 6337 (2017), 508–513.

[13] Rong Yang, Fernando Ordonez, and Milind Tambe. 2012. Computing optimal strategy against quantal response in security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. 847–854.