Carbon Market Simulation with Adaptive Mechanism Design

Han Wang[1] , **Wenhao Li**[1] , **Hongyuan Zha**[1] and **Baoxiang Wang**[1,2*]
[1]The Chinese University of Hong Kong, Shenzhen
[2]Vector Institute
xwanghan@gmail.com, {liwenhao,zhahy,bxiangwang}@cuhk.edu.cn

Abstract

A carbon market is a market-based tool that incentivizes economic agents to align individual profits with the global utility, i.e., reducing carbon emissions to tackle climate change. *Cap and trade* stands as a critical principle based on allocating and trading carbon allowances (carbon emission credit), enabling economic agents to follow planned emissions and penalizing excess emissions. A central authority is responsible for introducing and allocating those allowances in cap and trade. However, the complexity of carbon market dynamics makes accurate simulation intractable, which in turn hinders the design of effective allocation strategies. To address this, we propose an adaptive mechanism design framework, simulating the market using hierarchical, model-free multi-agent reinforcement learning (MARL). Government agents allocate carbon credits, while enterprises engage in economic activities and carbon trading. This framework illustrates agents' behavior comprehensively. Numerical results show MARL enables government agents to balance productivity, equality, and carbon emissions. Our project is available at https://github.com/xwanghan/Carbon-Simulator.

1 Introduction

Climate change has emerged as a pressing worldwide concern [Schmalensee *et al.*, 1998], significantly imperiling global ecosystems, economic systems, and sociopolitical stability. The United Nations reports that in developing regions, one in ten individuals subsists on less than US$ 1.90 daily [Nations, 2023b], with 2.2 billion people deprived of access to safely managed potable water resources [Nations, 2023c]. The burgeoning climate crisis amplifies these challenges, as worldwide temperature escalations provoke droughts and rising sea levels, exacerbating famines and enhanced forced displacements [Nations, 2023a].

In 2016, 196 nations endorsed the Paris Agreement to mitigate climate change collaboratively. However, pursuing transnational environmental targets often conflicts with short-term interests, requiring mechanisms to reconcile local and international objectives [UNFCCC, 2015]. Carbon markets exemplify such mechanisms [UNFCCC, 1997; Wara, 2007; Zhou and Li, 2019], incentivizing economic agents to curb emissions. The *cap and trade* format, predominant in carbon markets, involves *allocating* and *trading* allowances [Goulder and Schein, 2013; Schmalensee and Stavins, 2017; Zhou and Wang, 2016; Hepburn, 2007]. Economic agents must possess sufficient allowances to offset emissions or face penalties for surplus emissions. The *cap and trade* system sets a predetermined limit on allowances within an economy, with a central authority introducing and allocating allowances based on specified objectives. While this policy helps balance efficiency and fairness, determining the optimal allocation remains challenging in general economic contexts. The high-dimensional dynamics of the carbon market, influenced by rational, self-interested, and far-sighted economic agents, lead to market simulation reliance on models like CGE (computable general equilibrium) [Hübler *et al.*, 2014; Tang *et al.*, 2016; Bi *et al.*, 2019] or ABM (agent-based modeling) [Tang *et al.*, 2017; Zhou *et al.*, 2016; de Sousa, 2021] frameworks, employing simplifying assumptions that are arduous to validate, such as production and trading behaviors.

Given the unique nature of the carbon market, we integrate the *AI Economist* [Zheng *et al.*, 2022; Trott *et al.*, 2021] to simulate market dynamics. Our adaptive mechanism design framework, employing hierarchical, model-free MARL, mimics the carbon market. Lower-level enterprise agents engage in realistic economic activities, such as emitting carbon dioxide, trading emission credits, and investing in emission reduction projects. Higher-level government agents analyze diverse allocation strategies to achieve balanced efficiency and fairness, leading to significant carbon emission reductions. The framework demonstrates the conduct of rational, self-interested, and far-sighted agents within the carbon market. We emphasize that our approach is not a simple transfer of the AI-Economist from taxation to carbon credit allocation. Simulating the carbon market is challenging due to limited data, fluctuating regulations, and non-market factors.

To validate our simulator, we conducted comparisons with several widely adopted indicator allocation approaches at the firm level [Zhou and Wang, 2016]. The simulation results indicate reasonable action responses by enterprise agents to

*Corresponding to: Baoxiang Wang

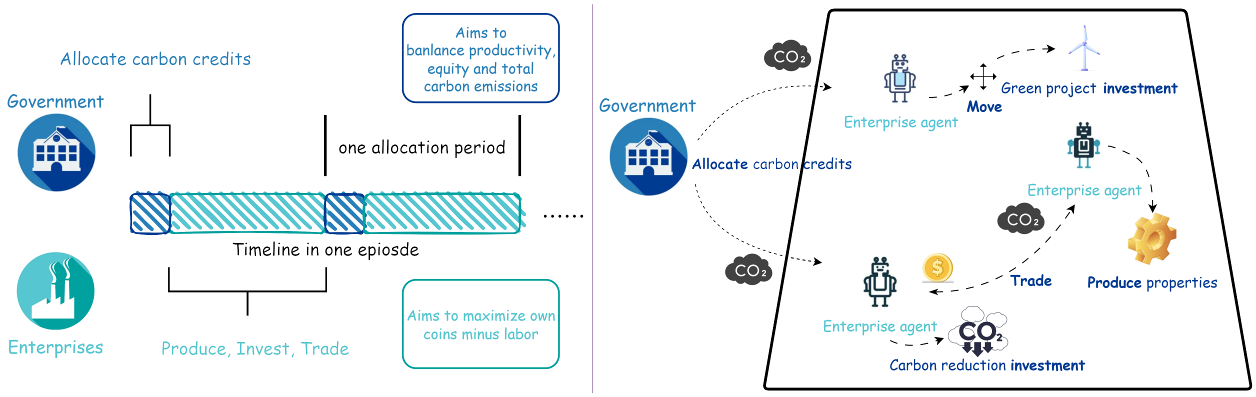

Figure 1: Simulator structure. **Left:** One episode is divided into several periods, and in each periods, the government firstly acts to allocate carbon credits; the remaining of time, enterprises do their economics activities. **Right:** Enterprises' economics activities are modeled in a Gather-Trade-Build game; in this grid map, they can produce properties (build) to get coins, can move which can gather carbon credits and increase community's total power (green project investment), do carbon reduction invest action to reduce the carbon emission level (carbon reduction investment), also trade carbon credits and coins with each other.

these allocation policies. Additionally, numerical findings demonstrate that government agents, through MARL, effectively discover allocation policies capable of balancing productivity, equity, and carbon emissions. Our primary contributions encompass: **1)** We propose a systematic carbon market simulator featuring carbon credits allocation and trading, and achieve realistic carbon economy simulation based on hierarchical, model-free MARL. **2)** We implement several widely adopted indicator allocation approaches at the firm level as baselines. **3)** We observe that learning-based allocation policies possess the potential to effectively balance productivity, equity, and carbon emissions.

2 Carbon Market Modeling

We present a carbon market framework which exhibits a hierarchical structure, consisting of the higher-level government RL agent and lower-level enterprise RL agents. Consequently, it is referred to as a hierarchical model-free MARL framework, which has also been known as a manager-worker architecture [Shu and Tian, 2019; Ma and Wu, 2020]. Concretely, the issues encountered by higher-level government agents can be modeled as a standard Markov decision process [Bellman, 1957] $\mathcal{M}_h = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$. The problems encountered by the lower-level enterprise agents can formally defined by a partially observable stochastic game [Hansen *et al.*, 2004] $\mathcal{G}_l = (\mathcal{S}, \mathcal{A}_{i=1}^{|\mathcal{I}|}, \mathcal{R}_{i=1}^{|\mathcal{I}|}, \mathcal{P}, \mathcal{O}_{i=1}^{|\mathcal{I}|}, \mathcal{I})$. And both government and enterprise are aimed to obtain optimal policies π that maximize their expected return $\mathbb{E}_{a^i \sim \pi_i, \mathbf{a}^{-i} \sim \pi_{-i}, s' \sim \mathcal{P}} \left[\sum_t \gamma^t r_t^i\right]$.

2.1 Model Details

Observation space $\mathcal{O}$. In real world, enterprise can observe their own attributes s (skills: enterprise size, R&D capabilities), assets x (income x^i, carbon emission level x^l and carbon emission credit x^c), and acquire request data in the carbon market from the news. While the government can collect data from all enterprises and access current market prices W_t. Additionally, because we model the economics activities based on a Gather-Trade-Build game [Zheng *et al.*, 2020], which effectively models general market-based behaviors, the position information O should also add to both observation. Therefore, the government's observation space is $\mathcal{O}_g = \{W_t, s, x, O\}$, while the enterprise agent i can observe $\mathcal{O}_i = \{s_i, x_i, O_i\}, i \in \{1, ..., \mathcal{I}\}$.

Action space $\mathcal{A}$. The government and enterprises are make decisions in different timescales (Figure 1). When enterprise agents act, they have the options to `Produce`, `Invest`, `Trade` and `Move`. Moveover, carbon emission credits are virtual resources which can be overdrafted. At the end of each year, when an enterprise settles carbon emission credits overdraft, the overdrafted portion will be penalized according to a unit price of p. The action of the government agent is a $\mathcal{I}+1$ dimensional discrete vector. At the beginning of each period, the government decides the allocation of credits for each enterprise as well as the total credits for the period.

Reward function $\mathcal{R}$. Each enterprise's objective is to maximize utility which is defined as [Debreu and others, 1954]:

$$r^i = (z^{1-\eta} - 1)/(1-\eta) - c^l * l, \qquad \eta = 0.23, \tag{1}$$

where z is income that the sum of enterprise's coins, c^l is labor's weight, and l is also the cumulative labor during all previous time steps. Inspired by [Zheng *et al.*, 2020], the government's objective is to maximize social welfare, which is defines as multiplication of productivity, equality and attenuation coefficient of excess carbon emissions:

$$r = \sum_i (x_i^c) * Gini(x^c) * exp(-c^e * ee), \tag{2}$$

where c^e is a coefficient and ee is total excess emissions.

2.2 Model Calibration

The structural parameters of simulation were calibrated to meet the following objectives: **1)** We calibrate the model of carbon emission reduction investments to possess similar attributes of risk, delay, and other behaviors observed in real-life investment activities [Lee, 2020]. Additionally, we calibrate parameters to ensure that the carbon price changes caused by carbon emission reduction investments are aligned with price tiers set within the carbon trading market. **2)** The penalty for excess carbon emissions is also calibrated such that the excess penalty is only offset when a certain level of carbon emission reduction investment by enterprises, alongside green project investments by all enterprises, is achieved. **3)** Furthermore, ensure that the outcomes resulting from various behaviors of enterprise agents are on the same temporal and economic scale.

3 Simulation and Visualization

Upon the completion of carbon market modeling, we can employ MARL to train the government (enterprise) agent(s), enabling them to exhibit behavior that closely resembles real-world scenarios, thereby achieving a realistic simulation of the carbon market.

3.1 Baseline Allocation Policies

We utilize the indicator approach [Zhou and Wang, 2016] to allocate the proportion of total carbon credits for each enterprise annually, with Emission (also called **grandfathering** or **GF** [Zetterberg *et al.*, 2012]), Emission intensity (also called **benchmarking** or **BM** [Groenenberg and Blok, 2002]), and Enterprise size selected as indicators (shorten as **SI**). Due to the large temporal scale of our simulation spanning 10 years, the government agent needs to allocate the total carbon emission credits for each year over the 10-year period. We refer to the global carbon emission historical data and future forecasts provided by the IPCC [Climate, 2023] to establish emission scenarios on a large temporal scale: we call this scenario **Convex**. Additionally, we provide scenarios where the annual emission targets decrease gradually over time (**Decreasing**), and scenarios where the annual emission targets remain constant over time (**Flat**).

3.2 MARL Allocation Policy Training

MARL training aims to discover an allocation policy that maximizes the Government's reward r, while also finding a balance between productivity, equality, and carbon emissions under the designated economy-climate coefficient c^e.

For the joint optimization of enterprise and government policies, we first initialize the parameters of enterprise agents to those trained under government policies based on **Flat** and Enterprise size as the indicator (**SI**) scenario. Subsequently, the parameters of government policies are randomly initialized. During training, we utilize the PPO algorithm [de Witt *et al.*, 2020] under the RLlib framework [Liang *et al.*, 2018]. Additionally, we experiment with various hyperparameters for both enterprise and government agents, including learning rate and entropy regularization. Following training with 400 million samples, we find both enterprise and government agents to converge to stable policies, which effectively balance productivity, equity, and carbon emissions. (Figure 2)

3.3 Visualization

We have also implemented visualization for the simulator, which provides a detailed breakdown of each time step within an episode. Through this visualization tool, we can gain a deeper understanding of and intuitive comparison between different baselines and MARL strategies (Figure 3).

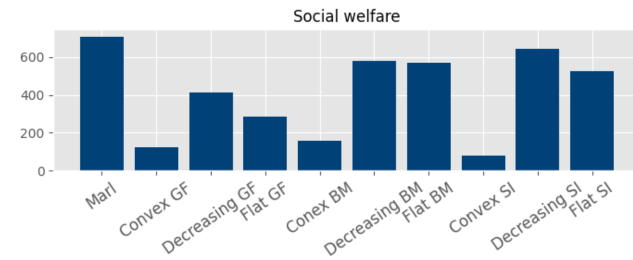

Figure 2: Quantitative results of different allocation policies.

Figure 3: Simulator dashboard, it presents detailed information encompassing enterprises' attributes, assets, and actions within a single time step across various example episodes under different policies. Additionally, it provides visual representations of the average carbon prices over different periods and presents rewards for both enterprises and government.

4 Closing Remarks

In this paper, enterprise and government agents participate in carbon market simulations via MARL-based adaptive mechanism design. By fine-tuning the government's reward function, we can exploit the adaptability to strike a balance between various economic and climate objectives. Unlike the commonly used indicator approach, MARL-based agents can incorporate more comprehensive information, enabling them to formulate more personalized and diversified allocation strategies. We also illustrates the practicality of employing hierarchical model-free MARL for carbon market simulation. It envisions the potential of machine learning to contribute to global emission reduction endeavors. However, the proposed simulator still needs to be improved, notably the absence of empirical modeling for emissions reduction investments made by enterprises. Consequently, future simulations can enhance their realism by integrating more real-world data.

Ethical Statement

In our development of the carbon market simulator, we adhere to principles of transparency, integrity, and fairness, ensuring compliance with the highest ethical standards while advancing understanding in environmental economics. We prioritize privacy, equity, and social responsibility throughout our research and development process. However, it's important to acknowledge that our simulator may not encompass all aspects of the real world. As such, we do not endorse the use of learned policies for actual policy making.

Acknowledgements

Baoxiang Wang is partially supported by National Natural Science Foundation of China (62106213, 72394361).

References

[Bellman, 1957] Richard Bellman. A markovian decision process. *Journal of mathematics and mechanics*, pages 679–684, 1957.

[Bi *et al.*, 2019] Huimin Bi, Hao Xiao, and Kejuan Sun. The impact of carbon market and carbon tax on green growth pathway in china: a dynamic cge model approach. *Emerging Markets Finance and Trade*, 55(6):1312–1325, 2019.

[Climate, 2023] Climate. *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, 2023.

[de Sousa, 2021] João Bernardo Narciso de Sousa. Carbon market multi-agent simulation metamodel: a testbed for coordination mechanisms in sociotechnical systems. 2021.

[de Witt *et al.*, 2020] Christian Schroeder de Witt, Tarun Gupta, Denys Makoviichuk, Viktor Makoviychuk, Philip HS Torr, Mingfei Sun, and Shimon Whiteson. Is independent learning all you need in the starcraft multi-agent challenge? *arXiv preprint arXiv:2011.09533*, 2020.

[Debreu and others, 1954] Gerard Debreu et al. Representation of a preference ordering by a numerical function. *Decision processes*, 3:159–165, 1954.

[Goulder and Schein, 2013] Lawrence H Goulder and Andrew R Schein. Carbon taxes versus cap and trade: a critical review. *Climate Change Economics*, 4(03):1350010, 2013.

[Groenenberg and Blok, 2002] Heleen Groenenberg and Kornelis Blok. Benchmark-based emission allocation in a cap-and-trade system. *Climate Policy*, 2(1):105–109, 2002.

[Hansen *et al.*, 2004] Eric A Hansen, Daniel S Bernstein, and Shlomo Zilberstein. Dynamic programming for partially observable stochastic games. In *AAAI*, 2004.

[Hepburn, 2007] Cameron Hepburn. Carbon trading: A review of the kyoto mechanisms. *Annual Review of Environment and Resources*, 32(1):375–393, 2007.

[Hübler *et al.*, 2014] Michael Hübler, Sebastian Voigt, and Andreas Löschel. Designing an emissions trading scheme for china—an up-to-date climate policy assessment. *Energy policy*, 75:57–72, 2014.

[Lee, 2020] Jun-Yeon Lee. Investing in carbon emissions reduction in the eoq model. *Journal of the Operational Research Society*, 71(8):1289–1300, 2020.

[Liang *et al.*, 2018] Eric Liang, Richard Liaw, Robert Nishihara, Philipp Moritz, Roy Fox, Ken Goldberg, Joseph Gonzalez, Michael Jordan, and Ion Stoica. Rllib: Abstractions for distributed reinforcement learning. In *International conference on machine learning*, pages 3053–3062. PMLR, 2018.

[Ma and Wu, 2020] Jinming Ma and Feng Wu. Feudal multi-agent deep reinforcement learning for traffic signal control. In *AAMAS*, 2020.

[Nations, 2023a] United Nations. Climate Change. https://www.un.org/en/global-issues/climate-change/, 2023. Accessed on May 30, 2024.

[Nations, 2023b] United Nations. Ending Poverty. https://www.un.org/en/global-issues/ending-poverty/, 2023. Accessed on May 30, 2024.

[Nations, 2023c] United Nations. Water. https://www.un.org/en/global-issues/water/, 2023. Accessed on May 30, 2024.

[Schmalensee and Stavins, 2017] Richard Schmalensee and Robert N Stavins. The design of environmental markets: What have we learned from experience with cap and trade? *Oxford Review of Economic Policy*, 33(4):572–588, 2017.

[Schmalensee *et al.*, 1998] Richard Schmalensee, Thomas M. Stoker, and Ruth A. Judson. World carbon dioxide emissions: 1950-2050. *The Review of Economics and Statistics*, 80(1):15–27, 1998.

[Shu and Tian, 2019] Tianmin Shu and Yuandong Tian. M^3RL: Mind-aware multi-agent management reinforcement learning. In *ICLR*, 2019.

[Tang *et al.*, 2016] Ling Tang, Jiarui Shi, and Qin Bao. Designing an emissions trading scheme for china with a dynamic computable general equilibrium model. *Energy Policy*, 97:507–520, 2016.

[Tang *et al.*, 2017] Ling Tang, Jiaqian Wu, Lean Yu, and Qin Bao. Carbon allowance auction design of china's emissions trading scheme: A multi-agent-based approach. *Energy Policy*, 102:30–40, 2017.

[Trott *et al.*, 2021] Alexander Trott, Sunil Srinivasa, Douwe van der Wal, Sebastien Haneuse, and Stephan Zheng. Building a foundation for data-driven, interpretable, and robust policy design using the ai economist. *arXiv preprint arXiv:2108.02904*, 2021.

[UNFCCC, 1997] UNFCCC. Kyoto protocol. 1997. Accessed on May 30, 2024.

[UNFCCC, 2015] UNFCCC. The paris agreement, 2015.

[Wara, 2007] Michael Wara. Is the global carbon market working? *Nature*, 445(7128):595–596, 2007.

[Zetterberg *et al.*, 2012] Lars Zetterberg, Markus Wråke, Thomas Sterner, Carolyn Fischer, and Dallas Burtraw. Short-run allocation of emissions allowances and long-term goals for climate policy. *Ambio*, 41:23–32, 2012.

[Zheng *et al.*, 2020] Stephan Zheng, Alexander Trott, Sunil Srinivasa, Nikhil Naik, Melvin Gruesbeck, David C Parkes, and Richard Socher. The ai economist: Improving equality and productivity with ai-driven tax policies. *arXiv preprint arXiv:2004.13332*, 2020.

[Zheng *et al.*, 2022] Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C Parkes, and Richard Socher. The ai economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science advances*, 8(18):eabk2607, 2022.

[Zhou and Li, 2019] Kaile Zhou and Yiwen Li. Carbon finance and carbon market in china: Progress and challenges. *Journal of Cleaner Production*, 214:536–549, 2019.

[Zhou and Wang, 2016] P. Zhou and M. Wang. Carbon dioxide emissions allocation: A review. *Ecological Economics*, 125:47–59, 2016.

[Zhou *et al.*, 2016] Meirong Zhou, Ming Zhou, Yanchun Pan, Zhimin Chen, and Jun Zeng. Multi-agent-based simulation for policy evaluation of carbon emissions. In *Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems: 16th Asia Simulation Conference and SCS Autumn Simulation Multi-Conference, AsiaSim/SCS AutumnSim 2016, Beijing, China, October 8-11, 2016, Proceedings, Part III 16*, pages 265–272. Springer, 2016.