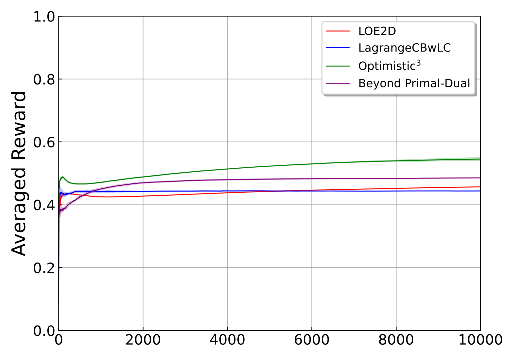
Additional experiment on BwK in both worlds

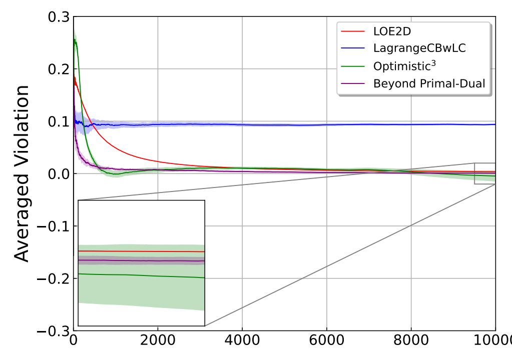
We evaluate our approach through numerical experiments on non-contextual multi-armed bandits across both stochastic and adversarial environments. We consider $|\mathcal{A}| = 10$ arms and the time horizon $T = 10000$. All results are obtained by averaging over 50 trials and reported with a 95% confidence interval.

- In the stochastic setting, each arm's mean reward $r(a) = \mu_a$ is drawn from the uniform distribution of $[0, 1]$ while mean constraint $c(a) = \lambda_a$ is drawn from the uniform distribution of $[-0.5, 1]$. Observations of rewards and costs are perturbed by Gaussian noise $\mathcal{N}(0, 0.05)$.
- In the adversarial setting, rewards and constraints adopt time-varying dynamics: $r_t(a) = \mu_a + \alpha_a^1 \sin(\omega_a^1 t)$ and $c_t(a) = \lambda_a + \alpha_a^2 \sin(\omega_a^2 t)$, where μ_a, λ_a match the stochastic setting, and frequencies $\omega_a^1, \omega_a^2 \sim \mathrm{Uniform}[0, 0.2]$. Observations of rewards and costs are perturbed by Gaussian noise $\mathcal{N}(0, 0.05)$.

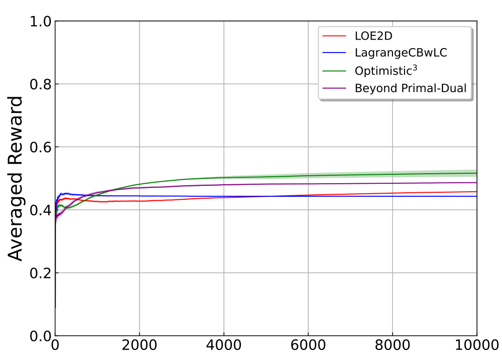
Figures 1 demonstrates that `Optimistic`^{3} outperforms all baseline algorithm, which justify our theoretical guarantees.



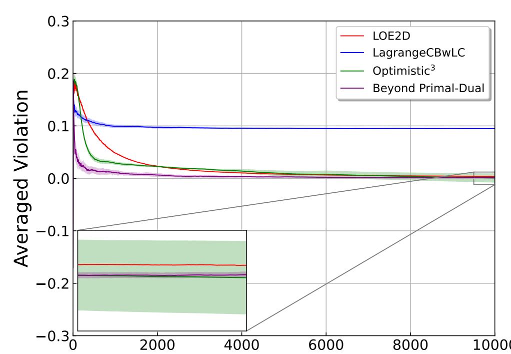
(a) Averaged reward under stochastic setting



(b) Averaged constraint violation under stochastic setting



(c) Averaged reward under adversarial setting



(d) Averaged constraint violation under adversarial setting

Figure 1. Averaged reward and constraint violation under LOE2D (Guo & Liu, 2024), LagrangeCBwLC (Slivkins et al., 2023), Beyond Primal-Dual (Bernasconi et al., 2024) and `Optimistic`^{3}.

References

Bernasconi, M., Castiglioni, M., Celli, A., and Fusco, F. Beyond primal-dual methods in bandits with stochastic and adversarial constraints. *arXiv preprint arXiv:2405.16118*, 2024.

Guo, H. and Liu, X. Stochastic constrained contextual bandits via lyapunov optimization based estimation to decision framework. In *The Thirty Seventh Annual Conference on Learning Theory*, pp. 2204–2231. PMLR, 2024.

Slivkins, A., Sankararaman, K. A., and Foster, D. J. Contextual bandits with packing and covering constraints: A modular lagrangian approach via regression. In *Annual Conference Computational Learning Theory*, 2023.