

## SEMINAR ANNOUNCEMENT

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING  
COLLEGE OF DESIGN AND ENGINEERING

Website: <https://cde.nus.edu.sg/ece>

**Area: Control, Intelligent Systems & Robotics (CISR)**

**Host: Prof Vincent Tan Yan Fu**

TOPIC	:	Maillard Sampling: Boltzmann Exploration Done Optimally
SPEAKER	:	Mr Bian Jie Graduate student, ECE Dept, NUS
DATE	:	Friday, 25 April 2025
TIME	:	10:00AM-11:00AM
VENUE	:	Join Zoom Meeting <a href="https://nus-sg.zoom.us/j/9839098564">https://nus-sg.zoom.us/j/9839098564</a>  Meeting ID: 9839098564 Passcode: no password

### ABSTRACT

The PhD thesis of Maillard (2013) presents a rather obscure algorithm for the K-armed bandit problem. This less-known algorithm, which we call Maillard sampling (MS), computes the probability of choosing each arm in a closed form, which is not true for Thompson sampling, a widely-adopted bandit algorithm in the industry. This means that the bandit-logged data from running MS can be readily used for counterfactual evaluation, unlike Thompson sampling. Motivated by such merit, we revisit MS and perform an improved analysis to show that it achieves both the asymptotical optimality and  $\sqrt{(KT \log T)}$  minimax regret bound where T is the time horizon, which matches the known bounds for asymptotically optimal UCB. We then propose a variant of MS called MS<sup>+</sup> that improves its minimax bound to  $\sqrt{(KT \log K)}$ . MS<sup>+</sup> can also be tuned to be aggressive (i.e., less exploration) without losing the asymptotic optimality, a unique feature unavailable from existing bandit algorithms. Our numerical evaluation shows the effectiveness of MS<sup>+</sup>.

### BIOGRAPHY

Bian Jie is currently pursuing his Ph.D. degree under the supervision of Professor Vincent Tan in the Department of ECE, NUS. His current research interests are in Multi-armed bandit.

<https://cde.nus.edu.sg/ece/highlights/events/>