# Bandits and Hyper-parameter Optimization: From Practice to Theory

**X. Shang**[1], R. de Reide[2], P. Ménard[1], E. Kaufmann[3], M. Valko[4]
[1] Inria Lille, SequeL team
[2] CWI
[3] CNRS, CRIStAL
[4] Google DeepMind

Inria SequeL

Lille, Feb. 2020

# What — Hyper-parameter optimization

## Problem

We tackle **hyper-parameter tuning** for *supervised learning* tasks:

# What — Hyper-parameter optimization

### Problem

We tackle **hyper-parameter tuning** for *supervised learning* tasks:

↳ global optimisation task: $\min\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Omega\}$;

# What — Hyper-parameter optimization

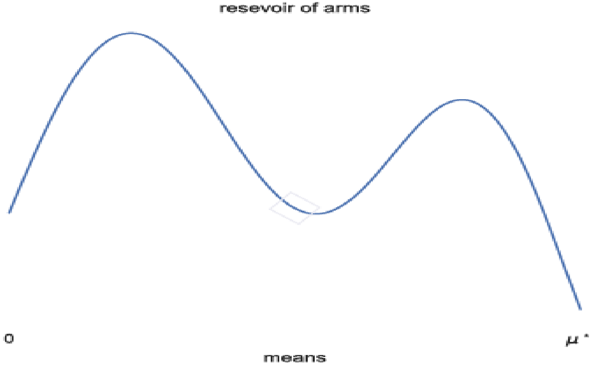
## Problem

We tackle **hyper-parameter tuning** for *supervised learning* tasks:

- ↳ global optimisation task: $\min\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Omega\}$;
- ↳ $f(\boldsymbol{\lambda}) \triangleq \mathbb{E}\left[\ell\left(\mathbf{Y}, \widehat{g}_{\boldsymbol{\lambda}}^{(n)}(\mathbf{X})\right)\right]$ measures the generalization power.

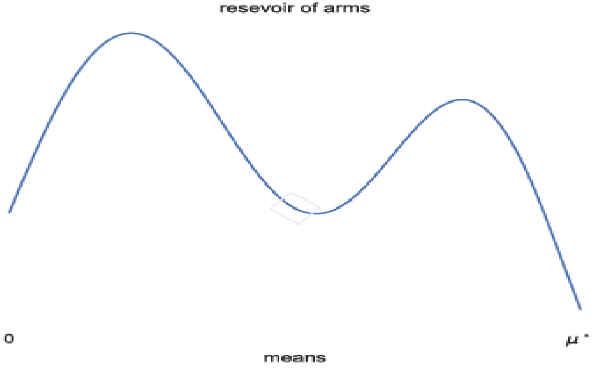
# How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* $\nu_0$.



# How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* $\nu_0$.



# How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* $\nu_0$.

In each round

1. (optional) query a new arm from $\nu_0$;
2. sample an arm that was previously queried.

goal: output an arm with mean close to $\mu^\star$
D-TTTS (Dynamic Top-Two Thompson Sampling) $\rightsquigarrow$ a dynamic algorithm built on TTTS.

# How — D-TTTS

## In this talk...

- Beta-Bernoulli Bayesian bandit model
- a uniform prior over the mean of new arms

**Initialization**: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$
**while** budget still available **do**
  $\forall i \in \mathcal{A}$, $\theta_i \sim \texttt{Beta}(S_i + 1, N_i - S_i + 1)$
  $I^{(1)} = \arg\max_{i=0,\dots,m} \theta_i$ {Thompson sampling}
**end while**

# How — D-TTTS

```
Initialization: μ_1 ~ ν_0; A = {μ_1}; m = 1; S_1, N_1 = 0
while budget still available do
    ∀i ∈ A, θ_i ~ Beta(S_i + 1, N_i − S_i + 1)
    I^(1) = arg max_{i=0,...,m} θ_i {Thompson sampling}
    if U(~ U([0, 1])) > β then
        while I^(2) ≠ I^(1) do
            ∀i ∈ A, θ'_i ~ Beta(S_i + 1, N_i − S_i + 1)
            I^(2) ← arg max_{i=0,...,m} θ'_i
        end while
        I^(1) ← I^(2)
    end if {TTTS}
end while
```

# How — D-TTTS

**Initialization**: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$

**while** budget still available **do**

$\quad \mu_{m+1} \sim \nu_0$; $\mathcal{A} \leftarrow \mathcal{A} \cup \{\mu_{m+1}\}$

$\quad S_{m+1}, N_{m+1} \leftarrow 0$; $m \leftarrow m + 1$

$\quad \forall i \in \mathcal{A},\ \theta_i \sim \mathtt{Beta}(S_i + 1, N_i - S_i + 1)$

$\quad I^{(1)} = \arg\max_{i=0,\dots,m} \theta_i$ {Thompson sampling}

$\quad$ **if** $U(\sim \mathcal{U}([0,1])) > \beta$ **then**

$\quad\quad$ **while** $I^{(2)} \neq I^{(1)}$ **do**

$\quad\quad\quad \forall i \in \mathcal{A}, \theta'_i \sim \mathtt{Beta}(S_i + 1, N_i - S_i + 1)$

$\quad\quad\quad I^{(2)} \leftarrow \arg\max_{i=0,\dots,m} \theta'_i$

$\quad\quad$ **end while**

$\quad\quad I^{(1)} \leftarrow I^{(2)}$

$\quad$ **end if** {TTTS}

$\quad Y \leftarrow \mathtt{evaluate\ arm}\ I^{(1)}$; $X \sim \mathtt{Ber}(Y)$

$\quad S_{I^{(1)}} \leftarrow S_{I^{(1)}} + X$; $N_{I^{(1)}} \leftarrow N_{I^{(1)}} + 1$

**end while**

# How — D-TTTS c'td

### D-TTTS in summary...

In each round, query a new arm endowed with a Beta(1,1) prior, without sampling it, and run TTTS on the new set of arms.

# How — D-TTTS c'td

## Order statistic trick

With $\mathcal{L}_{t-1}$ the list of arms that have been effectively sampled at time $t$, we run TTTS on the set $\mathcal{L}_{t-1} \cup \{\mu_0\}$ where $\mu_0$ is a pseudo-arm with posterior $\texttt{Beta}(t - |\mathcal{L}_{t-1}|, 1)$.

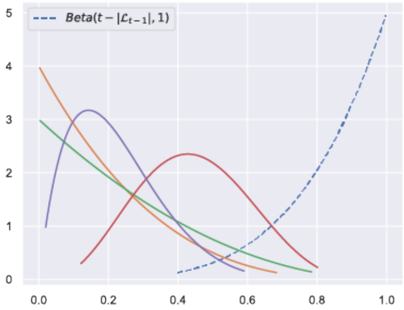



Figure: Posterior distributions of 4 arms and the pseudo-arm

# Why

1. `TTTS` is *anytime* for finitely-armed bandits;
2. the flexibility of this Bayesian algorithm allows to propose a **dynamic** version for the infinite BAI;
3. unlike previous approaches (for infinite-armed bandits or HPO), `D-TTTS` **does not** need to fix the number of arms queried in advance, and naturally **adapts** to the difficulty of the task.

# Why

1. `TTTS` is *anytime* for finitely-armed bandits;
2. the flexibility of this Bayesian algorithm allows to propose a **dynamic** version for the infinite BAI;

`SiRI` as an example for infinite-armed bandits

- choose a problem-dependent number of arms
- pull arms *optimistically*

# Why

1. `TTTS` is *anytime* for finitely-armed bandits;
2. the flexibility of this Bayesian algorithm allows to propose a **dynamic** version for the infinite BAI;

## `Hyperband` as an example for HPO

- equally divide the budget $B$ into $m$ brackets
- run sequential halving (Karnin et al. 2013) over a fixed number $n$ of configurations in each bracket
- trade-off between $B/mn$ and $n$

# HPO as a BAI problem

| BAI | HPO |
|---|---|
| query $\nu_0$ | pick a new configuration $\boldsymbol{\lambda}$ |
| sample an arm | train the classifier $g_{\boldsymbol{\lambda}}$ |
| reward | cross-validation loss |

# Experiments — Setting

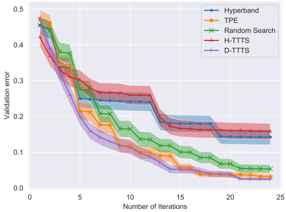
| Classifier | Hyper-parameter | Type | Bounds |
|---|---|---|---|
| SVM | $C$ | $\mathbb{R}^+$ | $\left[10^{-5}, 10^{5}\right]$ |
| | $\gamma$ | $\mathbb{R}^+$ | $\left[10^{-5}, 10^{5}\right]$ |

Table: hyper-parameters to be tuned for `UCI` experiments.

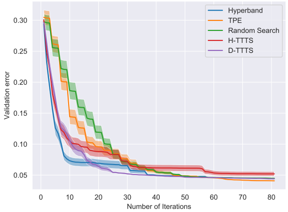
| Classifier | Hyper-parameter | Type | Bounds |
|---|---|---|---|
| MLP | hidden_layer_size | Integer | $[5, 50]$ |
| | alpha | $\mathbb{R}^+$ | $[0, 0.9]$ |
| | learning_rate_init | $\mathbb{R}^+$ | $\left[10^{-5}, 10^{-1}\right]$ |

Table: hyper-parameters to be tuned for `MNIST` experiments.

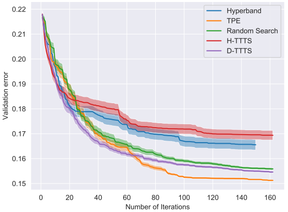
# Experiments — Some results



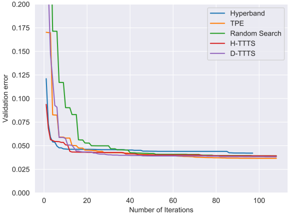

(a) wine




(b) breast cancer




(c) adult




(d) MNIST

# What now?

- Extend to the non-stochastic setting;

# What now?

- Extend to the non-stochastic setting;
- Theoretical guarantee?
  - Analysis of `D-TTTS` extremely difficult due to the dynamic nature;
  - How about take a step back to `TTTS` first?

**More details on `D-TTTS`**

Check out [Shang, Kaufmann, et al. 2019].

# BAI for finitely-armed bandits

- sampling rule;
- stopping rule;
- recommendation rule.

# What we know about TTTS... (Reminder)

**Input**: $\beta$
**for** $n = 1, 2, \ldots$ **do**
  $\forall i \in \mathcal{A}, \ \theta_i \sim \Pi_n$
  $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
  **if** $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
    **while** $I^{(2)} \neq I^{(1)}$ **do**
      $\forall i \in \mathcal{A}, \ \theta'_i \sim \Pi_n$
      $I^{(2)} \leftarrow \arg\max_{i=0,\ldots,m} \theta'_i$
    **end while**
    $I^{(1)} \leftarrow I^{(2)}$
  **end if**
  `evaluate arm` $I^{(1)}$
  `update` $\Pi_n$
**end for**

# What we know about TTTS... (Posterior convergence)

### Theorem (Russo 2016)

*Under TTTS, for **bounded** 1-dimensional exponential family priors, it holds almost surely that*

$$\lim_{n\to\infty} -\frac{1}{n}\log(1-\alpha_{n,I^\star}) = \Gamma^\star_\beta,$$

*where*

$$\alpha_{n,i} \triangleq \Pi_n(\theta_i > \max_{j\neq i} \theta_j).$$

# What we know about TTTS... (Complexity)

### Definition

Let $\Sigma_K = \{\boldsymbol{\omega} : \sum_{k=1}^K \omega_k = 1, \omega_k \geq 0\}$ and define for all $i \neq I^\star$

$$C_i(\omega, \omega') \triangleq \min_{x \in \mathcal{I}} \ \omega d(\mu_{I^\star}; x) + \omega' d(\mu_i; x),$$

where $d(\mu, \mu')$ is the KL-divergence. We define

$$\Gamma_\beta^\star \quad \triangleq \quad \max_{\substack{\boldsymbol{\omega} \in \Sigma_K \\ \omega_{I^\star} = \beta}} \min_{i \neq I^\star} C_i(\omega_{I^\star}, \omega_i).$$

# What we know about TTTS... (Complexity)

## Definition

Let $\Sigma_K = \{\boldsymbol{\omega} : \sum_{k=1}^{K} \omega_k = 1, \omega_k \geq 0\}$ and define for all $i \neq I^\star$

$$C_i(\omega, \omega') \triangleq \min_{x \in \mathcal{I}} \; \omega d(\mu_{I^\star}; x) + \omega' d(\mu_i; x),$$

where $d(\mu, \mu')$ is the KL-divergence. We define

$$\Gamma_\beta^\star \quad \triangleq \quad \max_{\substack{\boldsymbol{\omega} \in \Sigma_K \\ \omega_{I^\star} = \beta}} \min_{i \neq I^\star} C_i(\omega_{I^\star}, \omega_i).$$

In particular, for Gaussian bandits...

$$\Gamma_\beta^\star = \max_{\boldsymbol{\omega} : \omega_{I^\star} = \beta} \min_{i \neq I^\star} \frac{(\mu_{I^\star} - \mu_i)^2}{2\sigma^2 (1/\omega_i + 1/\beta)}.$$

# What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?

# What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?

# What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?

Lower bound

Under any $\delta$-correct strategy satisfying $T_{n,I^\star}/n \to \beta$,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\ln(1/\delta)} \geq \frac{1}{\Gamma^\star_\beta}.$$

# What we want to know about TTTS

- Can we remove the *boundedness* assumption of the prior?
- What can we say about the sample complexity in the fixed-confidence setting?
- Can we have finite-time guarantees?

# Main result — Posterior convergence

**Theorem (Shang, Heide, et al. 2020)**

*Under TTTS, for Gaussian bandits with improper Gaussian priors, it holds almost surely that*

$$\lim_{n \to \infty} -\frac{1}{n} \log(1 - \alpha_{n,I^\star}) = \Gamma_\beta^\star.$$

**Theorem (Shang, Heide, et al. 2020)**

*Under TTTS, for Bernoulli bandits and uniform priors, it holds almost surely that*

$$\lim_{n \to \infty} -\frac{1}{n} \log(1 - \alpha_{n,I^\star}) = \Gamma_\beta^\star.$$

# Main result — Sample complexity

**Theorem (Shang, Heide, et al. 2020)**

*The TTTS sampling rule coupled with the Chernoff stopping rule form a $\delta$-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies*

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

# Main result — Sample complexity

### Theorem (Shang, Heide, et al. 2020)

*The* TTTS *sampling rule coupled with the Chernoff stopping rule form a $\delta$-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies*

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

### Recall (Lower bound)

$$\liminf_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\ln(1/\delta)} \geq \frac{1}{\Gamma_\beta^\star}.$$

# Sample complexity sketch - $\delta$-correctness

Stopping rule

$$\tau_\delta^{\mathrm{Ch.}} \triangleq \inf \left\{ n \in \mathbb{N} : \max_{i \in \mathcal{A}} \min_{j \in \mathcal{A} \setminus \{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

# Sample complexity sketch - $\delta$-correctness

## Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf \left\{ n \in \mathbb{N} : \max_{i \in \mathcal{A}} \min_{j \in \mathcal{A} \setminus \{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

## Transportation cost

Let $\mu_{n,i,j} \triangleq (T_{n,i}\mu_{n,i} + T_{n,j}\mu_{n,j})/(T_{n,i} + T_{n,j})$, then we define

$$W_n(i,j) \triangleq \begin{cases} 0 & \text{if } \mu_{n,j} \geq \mu_{n,i}, \\ W_{n,i,j} + W_{n,j,i} & \text{otherwise,} \end{cases} \tag{2}$$

where $W_{n,i,j} \triangleq T_{n,i} d\left(\mu_{n,i}, \mu_{n,i,j}\right)$ for any $i, j$.

# Sample complexity sketch - $\delta$-correctness

## Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf \left\{ n \in \mathbb{N} : \max_{i \in \mathcal{A}} \min_{j \in \mathcal{A} \setminus \{i\}} W_n(i,j) > d_{n,\delta} \right\}. \tag{1}$$

In particular, for Gaussian bandits...

$$W_n(i,j) = \frac{(\mu_{n,i} - \mu_{n,j})^2}{2\sigma^2(1/T_{n,i} + 1/T_{n,j})} \mathbb{1}\{\mu_{n,j} < \mu_{n,i}\}.$$

# Sample complexity sketch - $\delta$-correctness

## Stopping rule

$$\tau_\delta^{\text{Ch.}} \triangleq \inf \left\{ n \in \mathbb{N} : \max_{i \in \mathcal{A}} \min_{j \in \mathcal{A} \setminus \{i\}} W_n(i,j) > d_{n,\delta} \right\}. \qquad (1)$$

## Theorem (Shang, Heide, et al. 2020)

*The TTTS sampling rule coupled with the Chernoff stopping rule* (1) *with a threshold $d_{n,\delta}$ derived from Kaufmann and Koolen 2018 and the recommendation rule $J_t = \arg\max_i \mu_{n,i}$, form a $\delta$-correct BAI strategy.*

# Sample complexity sketch — Sufficient condition for $\beta$-optimality

## Lemma

*Let $\delta, \beta \in (0,1)$. For any sampling rule which satisfies $\mathbb{E}\left[T_{\beta}^{\varepsilon}\right] < \infty$ for all $\varepsilon > 0$, we have*

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_{\delta}\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_{\beta}^{\star}},$$

*if the sampling rule is coupled with stopping rule (1).*

# Sample complexity sketch — Sufficient condition for $\beta$-optimality

**Lemma**

*Let $\delta, \beta \in (0,1)$. For any sampling rule which satisfies $\mathbb{E}\left[T_\beta^\varepsilon\right] < \infty$ for all $\varepsilon > 0$, we have*

$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star},$$

*if the sampling rule is coupled with stopping rule (1).*

$$T_\beta^\varepsilon \triangleq \inf\left\{N \in \mathbb{N} : \max_{i\in\mathcal{A}} |T_{n,i}/n - \omega_i^\beta| \leq \varepsilon, \forall n \geq N\right\}$$

# Sample complexity sketch - Core theorem

Theorem (Shang, Heide, et al. 2020)

*Under TTTS,* $\mathbb{E}\left[T_{\beta}^{\varepsilon}\right] < +\infty$.

The proof is inspired by Qin et al. (2017), but some technical novelties are introduced. In particular, our proof is much more intricate due to the randomized nature of the two candidate arms...

# Alleviate the computational burden?

**Input**: $\beta$
**for** $n = 1, 2, \dots$ **do**
  $\forall i \in \mathcal{A},\ \theta_i \sim \Pi_n$
  $I^{(1)} = \arg\max_{i=0,\dots,m} \theta_i$
  **if** $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
    **while** $I^{(2)} \neq I^{(1)}$ **do**
      $\forall i \in \mathcal{A},\ \theta'_i \sim \Pi_n$
      $I^{(2)} \leftarrow \arg\max_{i=0,\dots,m} \theta'_i$ {Re-sampling phase}
    **end while**
    $I^{(1)} \leftarrow I^{(2)}$
  **end if**
  `evaluate arm` $I^{(1)}$
  `update` $\Pi_n$
**end for**

# Alleviate the computational burden?

**Input**: $\beta$
**for** $n = 1, 2, \ldots$ **do**
    $\forall i \in \mathcal{A},\ \theta_i \sim \Pi_n$
    $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
    **if** $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
        $I^{(2)} \leftarrow \arg\min_{i \neq I^{(1)}} W_n(I^{(1)}, i)$ {T3C}
        $I^{(1)} \leftarrow I^{(2)}$
    **end if**
    evaluate arm $I^{(1)}$
    update $\Pi_n$
**end for**

# Main result — Sample complexity T3C

### Theorem (Shang, Heide, et al. 2020)

*The T3C sampling rule coupled with the Chernoff stopping rule form a $\delta$-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies*

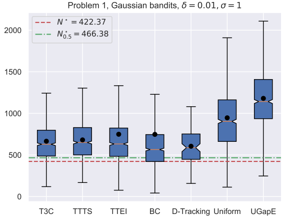
$$\limsup_{\delta \to 0} \frac{\mathbb{E}\left[\tau_\delta\right]}{\log(1/\delta)} \leq \frac{1}{\Gamma_\beta^\star}.$$

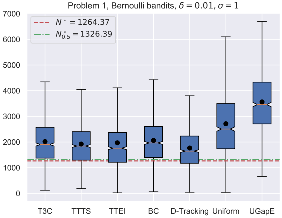
# Some illustrations — Time consumption

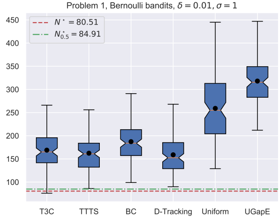
| **Sampling rule** | `T3C` | `TTTS` | `Uniform` |
|---|---|---|---|
| **Execution time (s)** | $1.6 \times 10^{-5}$ | $2.3 \times 10^{-4}$ | $6 \times 10^{-6}$ |

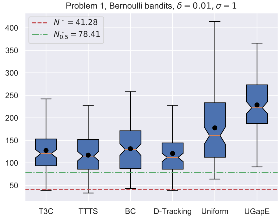
Table: average execution time in seconds for different sampling rules.

# Some illustrations — Average stopping time



# Still far from the Holy Grail...

- Finite-time analysis (fixed-budget setting?)

# Conclusion

More details on `TTTS` and `T3C`

Check out [Shang, Heide, et al. 2020].

# References

Thank you!


Zohar Karnin, Tomer Koren, and Oren Somekh. “Almost optimal exploration in multi-armed bandits”. In: *30th ICML*. 2013.

Emilie Kaufmann and Wouter Koolen. “Mixture martingales revisited with applications to sequential tests and confidence intervals”. In: 2018.

Daniel Russo. “Simple Bayesian algorithms for best arm identification”. In: *29th CoLT*. 2016.

Xuedong Shang, Rianne de Heide, Pierre Ménard, Emilie Kaufmann, and Michal Valko. “Fixed-confidence guarantees for Bayesian best-arm identification”. In: *23rd AIStats*. 2020.

Xuedong Shang, Emilie Kaufmann, and Michal Valko. “A simple dynamic bandit algorithm for hyper-parameter optimization”. In: *6th ICML Workshop on AutoML*. 2019.