MATHEMATICAL ENGINEERING TECHNICAL REPORTS

Stochastic Bandit Based on Empirical Moments

Junya HONDA and Akimichi TAKEMURA

Stochastic Bandit Based on Empirical Moments

Junya Honda[1★] and Akimichi Takemura[2★★]

[1] Graduate school of Frontier Sciences,
[2] Graduate school of Information Science and Technology,
The University of Tokyo, Japan.
{honda,takemura}@stat.t.u-tokyo.ac.jp

Abstract. In the multiarmed bandit problem a gambler chooses an arm of a slot machine to pull considering a tradeoff between exploration and exploitation. We study the stochastic bandit problem where each arm has a reward distribution supported in a known bounded interval, e.g. $[0, 1]$. For this model, policies which take into account the empirical variances (i.e. second moments) of the arms are known to perform effectively. In this paper, we generalize this idea and we propose a policy which exploits the first d empirical moments for arbitrary d fixed in advance. The asymptotic upper bound of the regret of the policy approaches the theoretical bound by Burnetas and Katehakis as d increases. By choosing appropriate d, the proposed policy realizes a tradeoff between the computational complexity and the expected regret.

1 Introduction

The multiarmed bandit problem is one of the formulations of the tradeoff between exploration and exploitation. This problem is based on an analogy with a gambler playing a slot machine with more than one arm. The gambler pulls arms sequentially so that the total reward is maximized.

We consider a K-armed stochastic bandit problem originally considered in [1]. There are K arms and each arm $i = 1, \cdots, K$ has a probability distribution F_i with the expected value μ_i. The gambler chooses an arm to pull based on a policy and receives a reward according to F_i independently in each round. We call an arm i optimal if $\mu_i = \mu^*$ and suboptimal if $\mu_i < \mu^*$. Then, the goal of the gambler is to maximize the sum of the rewards by pulling optimal arms as often as possible. Many researches have been conducted for the stochastic bandit problem [2, 3, 4, 5, 6, 7] as well as the non-stochastic bandit [8, 9].

In this paper we consider the model $\mathcal{F}$, the family of distributions with supports contained in the bounded interval $[0, 1]$. The gambler knows that each distribution F_i is included in $\mathcal{F}$. For this model Upper Confidence Bound (UCB) policies are popular for their simple form and fine performance [10, 11]. Recently Honda and Takemura [12] proposed Deterministic Minimum Empirical

Divergence (DMED) policy which satisfies for arbitrary suboptimal arm i that

$$\mathrm{E}[T_i(n)] \leq \left(\frac{1}{D_{\min}(F_i, \mu^*)} + \mathrm{o}(1) \right) \log n \tag{1}$$

where $T_i(n)$ denotes the number of times that arm i has been pulled over the first n round and

$$D_{\min}(F, \mu) \equiv \min_{G \in \mathcal{F}: \mathrm{E}_G[X] \geq \mu} D(F \| G)$$

with Kullback-Leibler divergence $D(\cdot \| \cdot)$. DMED is asymptotically optimal since the coefficient of $\log n$ on the right-hand side of (1) coincides with the theoretical bound given in [13]. However, the complexity of the DMED policy is still larger than e.g. UCB policies, although the computation involved in DMED is formulated as a univariate convex optimization. It is mainly because DMED requires the empirical distributions of the arms themselves whereas other popular policies can be computed by the moments of the empirical distributions of the arms, such as means and variances.

Now, our question is how we can bring the performance close to the right-hand side of (1) by a policy which only considers the first d empirical moments of the arms at each round. In this paper, we propose *DMED-M* policy which is a variant of DMED and is computable only by the empirical moments of the arms. For arbitrary suboptimal arm i, DMED-M satisfies

$$\mathrm{E}[T_i(n)] \leq \left(\frac{1}{\inf_{F \in \mathcal{F}: \mathrm{E}^{(d)}(F) = \mathrm{E}^{(d)}(F_i)} D_{\min}(F, \mu^*)} + \mathrm{o}(1) \right) \log n \ , \tag{2}$$

where $\mathrm{E}^{(d)}(F) \equiv (\mathrm{E}_{F_i}[X], \cdots, \mathrm{E}_{F_i}[X^d])$ denotes the first d moments of F and this upper bound approaches (1) as $d \to \infty$.

DMED-M is obtained by an analogy with DMED. Intuitively, DMED exploits the fact that the maximum likelihood that the arm with empirical distribution F_i is actually the best is roughly $\exp(-t D_{\min}(F_i, \mu^*))$ for number of samples t. When ignoring properties of the distribution F_i except for its first d moments, we overestimate the maximum likelihood as

$$\exp \left(- t \inf_{F \in \mathcal{F}: \mathrm{E}^{(d)}(F) = \mathrm{E}^{(d)}(F_i)} D_{\min}(F, \mu^*) \right)$$

instead of $\exp(-t D_{\min}(F_i, \mu^*))$ and the bound (2) appears correspondingly.

In DMED-M, it is necessary to compute $\inf_{F \in \mathcal{F}: \mathrm{E}^{(d)}(F) = (M_1, \cdots, M_d)} D_{\min}(F, \mu)$ for each round. Classical results on *Tchebysheff systems* and *moment spaces* reveal that $\bar{F}$ attaining the infimum is determined only by the value of the first d moments $(M_1, \cdots, M_d)$ when the objective function $D_{\min}(\cdot, \mu)$ is included in a particular class. Therefore the infimum is obtained by computing firstly the optimal solution $\bar{F}$ and then the value of the function $D_{\min}(\bar{F}, \mu)$. Both are obtained by solving polynomial equations and DMED-M can be computed efficiently for small d.

This paper is organized as follows. In Sect. 2, we give definitions used throughout this paper. We propose DMED-M policy in Sect. 3. In Sect. 4, we study the minimization of $D_{\min}$ over distributions whose first d moments are common for a practical implementation of DMED-M. Proofs of results in Sects. 3 and 4 are given in Sect. 5. In Sect. 6, we discuss an improvement of DMED-M in terms of the worst case performance. We present some simulation results on DMED-M in Sect. 7. We conclude the paper with some remarks in Sect. 8.

2 Preliminaries

Let $\mathcal{F}$ be the family of probability distributions on $[0, 1]$ and $F_i \in \mathcal{F}$ be the distribution of the arm $i = 1, \ldots, K$. $\mathrm{E}_F[\cdot]$ denotes the expectation under $F \in \mathcal{F}$. When we write e.g. $\mathrm{E}_F[u(X)]$ for a function $u : \mathbb{R} \to \mathbb{R}$, X denotes a random variable with distribution F. A set of probability distributions for K arms is denoted by $\boldsymbol{F} \equiv (F_1, \ldots, F_K) \in \mathcal{F}^K \equiv \prod_{i=1}^{K} \mathcal{F}$. The expected value of arm i is denoted by $\mu_i \equiv \mathrm{E}_{F_i}[X]$ and the optimal expected value is denoted by $\mu^* \equiv \max_i \mu_i$.

Let $T_i(n)$ be the number of times that arm i has been pulled through the first n rounds. $\hat{F}_i(n)$ and $\hat{\mu}_i(n)$ denote the empirical distribution and the mean of arm i after the first n rounds, respectively. $\hat{\mu}^*(n) \equiv \max_i \hat{\mu}_i(n)$ denotes the highest empirical mean after the first n rounds. We call an arm i a current best if $\hat{\mu}_i(n) = \hat{\mu}^*(n)$.

Now we review results in [12]. Define an index for $F \in \mathcal{F}$ and $\mu \in [0, 1]$

$$D_{\min}(F, \mu) \equiv \min_{G \in \mathcal{F}: \mathrm{E}(G) \geq \mu} D(F\|G) \ ,$$

where Kullback-Leibler divergence $D(F\|G)$ is given by

$$D(F\|G) \equiv \begin{cases} \mathrm{E}_F\left[\log \frac{\mathrm{d}F}{\mathrm{d}G}\right] & \frac{\mathrm{d}F}{\mathrm{d}G} \text{ exists,} \\ +\infty & \text{otherwise.} \end{cases}$$

Under DMED policy proposed in [12], the expectation of $T_i(n)$ for any suboptimal arm i is bounded as

$$\mathrm{E}_{\boldsymbol{F}}[T_i(n)] \leq \frac{1+\epsilon}{D_{\min}(F_i, \mu^*)} \log n + \mathrm{O}(1) \tag{3}$$

where $\epsilon > 0$ is arbitrary. The coefficient of the logarithmic term $1/D_{\min}(F_i, \mu^*)$ is the best possible [13] and the following property holds for the function $D_{\min}(F, \mu)$.

Proposition 1 ([12, Theorems 5 and 8]). *If* $\mathrm{E}_F[X] \geq \mu$ *then* $D_{\min}(F, \mu) = 0$. *If* $\mathrm{E}_F[X] < \mu = 1$ *then* $D_{\min}(F, \mu) = \infty$. *If* $\mathrm{E}_F[X] < \mu < 1$,

$$D_{\min}(F, \mu) = \max_{0 \leq \nu \leq \frac{1}{1-\mu}} \mathrm{E}_F[\log(1 - (X - \mu)\nu)]$$

$$= \begin{cases} \mathrm{E}_F\left[\log\left(1 - X\right)\right] - \log(1 - \mu) & \mathrm{E}_F\left[\frac{1}{1-X}\right] \leq \frac{1}{1-\mu}, \\ \max_{0 < \nu < \frac{1}{1-\mu}} \mathrm{E}_F[\log(1 - (X - \mu)\nu)] & \text{otherwise,} \end{cases}$$

where we define $\log 0 \equiv -\infty$.

Let $\mathrm{E}^{(d)}(F) \equiv (\mathrm{E}_F[X], \cdots, \mathrm{E}_F[X^d])$ denote the first d moments of F. The set of distributions with the first d moments equal to $\boldsymbol{M} = (M_1, \cdots, M_d)$ is defined as $\mathcal{F}(\boldsymbol{M}) \equiv \{F \in \mathcal{F} : \mathrm{E}^{(d)}(F) = \boldsymbol{M}\}$. We sometimes write $\boldsymbol{M}^{(d)}$ instead of $\boldsymbol{M}$ to clarify the length of the vector.

Now define $D_{\min}^{(d)}(\boldsymbol{M}, \mu)$ by

$$D_{\min}^{(d)}(\boldsymbol{M}, \mu) \equiv \inf_{F \in \mathcal{F}(\boldsymbol{M})} D_{\min}(F, \mu) \ .$$

This function $D_{\min}^{(d)}$ plays a central role throughout this paper.

3 DMED-M Policy

In this section we introduce DMED-M policy. This policy determines an arm to pull based on the empirical moments of the arms. DMED-M requires computation of the function $D_{\min}^{(d)}$ and we analyze this function in the next section.

In the following algorithm, each arm is pulled at most once in one loop. Through the loop, the list of arms pulled in the next loop is determined. L_C denotes the list of arms to be pulled in the current loop. L_N denotes the list of arms to be pulled in the next loop. $L_R \subset L_C$ denotes the list of remaining arms of L_C which have not yet been pulled in the current loop. The criterion for choosing an arm i is the occurrence of the event $J_i(n)$ given by

$$J_i(n) \equiv \{T_i(n) D_{\min}^{(d)}(\mathrm{E}^{(d)}(\hat{F}_i(n)), \hat{\mu}^*(n)) \leq \log n - \log T_i(n)\}, \qquad (4)$$

where $\mathrm{E}^{(d)}(\hat{F}_i(n))$ represents the first d empirical moments of arm i.

> **[DMED-M Policy]**
> **Parameter.** Integer $d > 0$.
> **Initialization.** $L_C, L_R := \{1, \cdots, K\}$, $L_N := \emptyset$. Pull each arm once. $n := K$.
> **Loop.**
> 1. For $i \in L_C$ in the ascending order,
>   1.1. $n := n + 1$ and pull arm i. $L_R := L_R \setminus \{i\}$.
>   1.2. $L_N := L_N \cup \{j\}$ (without a duplicate) for all $j \notin L_R$ such that $J_j(n)$ occurs.
> 2. $L_C, L_R := L_N$ and $L_N := \emptyset$.

As shown above, $|L_C|$ arms are pulled in one loop. At every round, arm i is added to L_N if $J_i(n)$ occurs unless $i \in L_R$, that is, arm i is planned to be pulled in the remaining rounds in the current loop. Note that if arm i is a current best for the n-th round then $J_i(n)$ holds since $D_{\min}(\mathrm{E}^{(m)}(\hat{F}_i(n)), \hat{\mu}^*(n)) = 0$ for this case. Then L_C is never empty. Note that DMED in [12] is obtained by replacing $D_{\min}^{(d)}(\mathrm{E}^{(d)}(\hat{F}_i(n)), \hat{\mu}^*(n))$ in (4) by $D_{\min}(\hat{F}_i(n), \mu)$. In view of Theorem 2 below, DMED can be regarded as DMED-M with $d = \infty$.

Theorem 1. *Fix $\boldsymbol{F} \in \mathcal{F}^K$ for which there exists a unique optimal arm j. Under DMED-M policy, for any suboptimal arm i and $\epsilon > 0$ it holds that*

$$\mathrm{E}_{\boldsymbol{F}}[T_i(n)] \leq \frac{1+\epsilon}{D_{\min}^{(d)}(\mathrm{E}^{(d)}(F_i), \mu^*)} \log n + \mathrm{O}(1)$$

where $\mathrm{O}(1)$ denotes a constant dependent on ϵ and $\boldsymbol{F}$ but independent of n.

This theorem can be proved in a similar way as Theorem 4 of [12] with the fact that $D_{\min}(F, \mu) \geq D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu)$ always holds. However, we omit the proof because it is long and very similar to the proof of Theorem 4 of [12]. The bound in Theorem 1 approaches that of DMED given by (3) as $d \to \infty$ from the following theorem, which we show in Sect. 5.

Theorem 2. *For arbitrary $F \in \mathcal{F}$ it holds that*

$$\lim_{d \to \infty} D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu) = D_{\min}(F, \mu) \ .$$

4 Practical Representation of $D_{\min}^{(d)}$

For a computation and a theoretical evaluation of DMED, it is essential to analyze the function $D_{\min}^{(d)}(\boldsymbol{M}, \mu) = \inf_{F \in \mathcal{F}:\mathrm{E}^{(d)}(F)=\boldsymbol{M}} D_{\min}(F, \mu)$. In this section we study an explicit representation of this function.

The following theorem is the main result of the paper. In this theorem, we identify a pair $(\{x_i\}, \{f_i\})$ with a discrete distribution such that $F(\{x_i\}) = f_i$.

Theorem 3. *If $\mathcal{F}(\boldsymbol{M}) = \{F \in \mathcal{F} : \mathrm{E}^{(d)}(F) = \boldsymbol{M}\}$ is nonempty then there exists a unique optimal solution $\bar{F} \in \mathcal{F}$ such that*

$$D_{\min}^{(d)}(\boldsymbol{M}, \mu) = \inf_{F \in \mathcal{F}(\boldsymbol{M})} D_{\min}(F, \mu) = D_{\min}(\bar{F}, \mu) \ . \tag{5}$$

Furthermore, $F \in \mathcal{F}$ is the unique optimal solution $\bar{F}$ if and only if $\big((x_1, \cdots, x_l), (f_1, \cdots, f_l)\big)$ for $l = \lceil d/2 \rceil + 1$ is a solution of

$$\begin{cases} \sum_{i=1}^{l} f_i x_i^m = M_m \ (m = 0, \cdots, d), \ x_1 = 0, \ x_l = 1, & d \text{ is odd}, \\ \sum_{i=1}^{l} f_i x_i^m = M_m \ (m = 0, \cdots, d), & x_l = 1, & d \text{ is even}, \end{cases} \tag{6}$$

where we define the zeroth moment as $M_0 = 1$.

Note that the above $\bar{F}$ only depends on the moment $\boldsymbol{M}$. Then, the value of $D_{\min}^{(d)}(\boldsymbol{M}, \mu)$ is obtained by computing $\bar{F}$ first and then $D_{\min}(\bar{F}, \mu)$. Recall that $D_{\min}(\bar{F}, \mu) = \max_{0 \leq \nu \leq (1-\mu)^{-1}} \mathrm{E}_{\bar{F}}[\log(1 - (X - \mu)\nu)]$. Since $\bar{F}$ has finite support $\{x_1, \cdots, x_l\}$, the optimal solution ν^* attaining the maximum is one of the boundary points 0, $(1-\mu)^{-1}$ or an interior point $\nu \in (0, (1-\mu)^{-1})$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\nu} \mathrm{E}_{\bar{F}}[\log(1 - (X - \mu)\nu)] = \frac{\sum_{i=1}^{l}(\mu - x_i)\prod_{j \neq i}(1 - (x_j - \mu)\nu)}{\prod_{i=1}^{l}(1 - (x_i - \mu)\nu)} = 0 \ , \tag{7}$$

which is obtained by solving the l-th degree polynomial equation. We give an explicit form of $D_{\min}^{(d)}(\boldsymbol{M}, \mu)$ for $d = 1, 2, 3$ in the following theorem.

Theorem 4. *If $M_1 < \mu < 1$ and $\mathcal{F}(\boldsymbol{M})$ is nonempty, then $D_{\min}^{(d)}(\boldsymbol{M}, \mu)$ is expressed for $d = 1, 2$ as*

$$D_{\min}^{(1)}(\boldsymbol{M}, \mu) = (1 - M_1) \log \frac{1 - M_1}{1 - \mu} + M_1 \log \frac{M_1}{\mu} \ ,$$

$$D_{\min}^{(2)}(\boldsymbol{M}, \mu) = \frac{(1 - M_1)^2}{1 - 2M_1 + M_2} \log \left(1 - \left(\frac{M_1 - M_2}{1 - M_1} - \mu \right) \nu^{(2)} \right)$$
$$+ \frac{M_2 - M_1^2}{1 - 2M_1 + M_2} \log \left(1 - (1 - \mu) \nu^{(2)} \right)$$

where

$$\nu^{(2)} = \frac{(1 - M_1)(M_1 - \mu)}{(1 - M_1)\mu^2 - (1 - M_2)\mu + M_1 - M_2} \ .$$

For $d = 3$ it is expressed as

$$D_{\min}^{(3)}(\boldsymbol{M}, \mu) = \begin{cases} D_{\min}^{(1)}(\boldsymbol{M}, \mu), & M_1 = M_2 = M_3, \\ \sum_{l=1}^{3} f_l \log(1 - \bar{x}_l \nu^{(3)}) & \text{otherwise}, \end{cases}$$

where

$$(\bar{x}_1, \bar{x}_2, \bar{x}_3) = \left(-\mu, \frac{M_2 - M_3}{M_1 - M_2} - \mu, 1 - \mu \right),$$

$$(f_2, f_3) = \left(\frac{(M_1 - M_2)^3}{(M_2 - M_3)(M_1 - 2M_2 + M_3)}, \frac{M_1 M_3 - M_2^2}{M_1 - 2M_2 + M_3} \right), f_1 = 1 - f_2 - f_3,$$

$$(8)$$

$$\nu^{(3)} = \begin{cases} \frac{-b + \sqrt{b^2 + 4ac}}{2a}, & a \neq 0, \\ \frac{c}{b}, & a = 0, \end{cases}$$

for

$$(a, b, c) = \left(\bar{x}_1 \bar{x}_2 \bar{x}_3, \ (M_2 - 2\mu M_1 + \mu^2) + (\bar{x}_1 + \bar{x}_2 + \bar{x}_3)(\mu - M_1), \ \mu - M_1 \right) \ .$$

This theorem is obtained by solving (7) with $\bar{F} = \bar{F}^{(d)}$ given in Lemma 1 below.

Lemma 1. *If $\mathcal{F}(\boldsymbol{M})$ is nonempty then the solution $\bar{F}^{(d)}$ of (6) is expressed for $d = 1, 2, 3$ as*

$$\bar{F}^{(1)} = (1 - M_1)\delta(0) + M_1 \delta(1),$$

$$\bar{F}^{(2)} = \begin{cases} \delta(1) & M_1 = M_2 = 1, \\ \frac{(1 - M_1)^2}{1 - 2M_1 + M_2} \delta\left(\frac{M_1 - M_2}{1 - M_1} \right) + \frac{M_2 - M_1^2}{1 - 2M_1 + M_2} \delta(1) & \text{otherwise}, \end{cases}$$

$$\bar{F}^{(3)} = \begin{cases} \bar{F}^{(1)} & M_1 = M_2 = M_3, \\ f_1 \delta(0) + f_2 \delta\left(\frac{M_2 - M_3}{M_1 - M_2} \right) + f_3 \delta(1) & \text{otherwise}, \end{cases}$$

where $\delta(x)$ denotes the delta measure at x and (f_1, f_2, f_3) is given by (8).

This lemma can be confirmed easily by substitution of $\bar{F}^{(d)}$ ($d = 1, 2, 3$) into (6).

5 Proofs

In this section we show Theorems 2 and 3.

5.1 A Proof of Theorem 2

Theorem 2 is proved by a basic result on weak convergence and Lévy distance (see, e.g., [14]). We say that a sequence of probability distributions $\{F_i\}$ converges weakly to F if $\lim_{i\to\infty} \mathrm{E}_{F_i}[u(X)] = \mathrm{E}_F[u(X)]$ for all bounded, continuous function $u(x)$. Define the Lévy distance $L(\cdot, \cdot)$ as

$$L(F, G) = \inf\{h > 0 : \forall x,\, F(x - h) - h \le G(x) \le F(x + h) + h\}$$

where $F(\cdot)$ and $G(\cdot)$ denote cumulative distribution functions. A weak convergence is equivalent to the convergence of the Lévy distance, that is, $\{F_i\}$ converges weakly to F if and only if $\lim_{i\to\infty} L(F_i, F) = 0$.

Proposition 2 ([12, Theorem 7]). $D_{\min}(F, \mu)$ *is continuous in $F \in \mathcal{F}$ with respect to the Lévy distance.*

Now we show Theorem 2 by Prop. 2.

Proof (of Theorem 2). From the continuity of $D_{\min}(F, \mu)$ in F, it suffices to show for $\boldsymbol{M}^{(d)} = \mathrm{E}^{(d)}(F)$ that

$$\limsup_{d\to\infty} \sup_{G \in \mathcal{F}(\boldsymbol{M}^{(d)})} L(G, F) = 0 \ . \tag{9}$$

Let $\{G_d \in \mathcal{F}(\boldsymbol{M}^{(d)})\}_{d=1,2,\cdots}$ be a sequence such that

$$\limsup_{d\to\infty} \sup_{G \in \mathcal{F}(\boldsymbol{M}^{(d)})} L(F_d, F) = \limsup_{d\to\infty} L(G_d, F) =: \bar{L} \ .$$

Since $\mathcal{F} \supset \mathcal{F}(\boldsymbol{M}^{(d)})$ is compact with respect to the Lévy distance, there exist $\bar{G} \in \mathcal{F}$ and a convergent subsequence $\{G_{d_i}\}$ of $\{G_d\}$ such that

$$\lim_{i\to\infty} L(G_{d_i}, F) = \bar{L} \ , \tag{10}$$

$$\lim_{i\to\infty} L(G_{d_i}, \bar{G}) = 0 \ , \tag{11}$$

where (11) means that $\{G_{d_i}\}$ converges weakly to $\bar{G}$. From the definition of weak convergence, for all natural numbers $m \in \mathbb{N}$ it holds that $\lim_{i\to\infty} \mathrm{E}_{G_{d_i}}[X^m] = \mathrm{E}_{\bar{G}}[X^m]$. On the other hand, $\mathrm{E}_{G_{d_i}}[X^m] = \mathrm{E}_F[X^m]$ for all $d_i \ge m$ from $G_{d_i} \in \mathcal{F}(\boldsymbol{M}^{(d_i)})$. Therefore we obtain for all $m \in \mathbb{N}$ that

$$\mathrm{E}_F[X^m] = \lim_{i\to\infty} \mathrm{E}_{G_{d_i}}[X^m] = \mathrm{E}_{\bar{G}}[X^m] \ .$$

Note that a sequence of moments $\{\mathrm{E}_F[X^m]\}$ has one-to-one correspondence to a distribution F for the case of bounded support. Therefore $\bar{G} = F$ and we obtain $\bar{L} = 0$ from (10). $\qquad\square$

5.2 A Proof of Theorem 3

Theorem 3 is proved by theories of Tchebysheff systems and moment spaces (see Appendix), and the basic result on a saddle-point in the following. For a function $\varphi(x, y) : \mathcal{X} \times \mathcal{Y} \to [-\infty, +\infty]$, a point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is called a saddle-point if $\varphi(\bar{x}, y) \leq \varphi(\bar{x}, \bar{y}) \leq \varphi(x, \bar{y})$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. A necessary and sufficient condition for a saddle-point is

$$\sup_{y \in \mathcal{Y}} \varphi(\bar{x}, y) = \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \varphi(x, y) = \sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \varphi(x, y) = \inf_{x \in \mathcal{X}} \varphi(x, \bar{y}) \ .$$

Proposition 3 (Minimax Theorem [15]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be a compact subset of a topological vector space $\mathcal{V}$ and $\mathcal{U}$. Let $\varphi(x, y) : \mathcal{X} \times \mathcal{Y} \to [-\infty, +\infty]$ be a function such that $\varphi(\cdot, y)$ is convex and lower-semicontinuous for any fixed x and $\varphi(x, \cdot)$ is concave and upper-semicontinuous for any fixed y. Then there exists a saddle point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$.*

In the proof of Theorem 3, we regard a probability measure F as an element of the family $\mathcal{V}$ of positive measures on $[0, 1]$ to exploit the results in Appendix. By letting $\tilde{\boldsymbol{M}} := (1, M_1, \cdots, M_d)$ for $\boldsymbol{M} = (M_1, \cdots, M_d)$, $D_{\min}^{(d)}(\boldsymbol{M}, \mu)$ is rewritten as

$$D_{\min}(\boldsymbol{M}, \mu) = \inf_{F \in \mathcal{V}(\tilde{\boldsymbol{M}})} \max_{0 \leq \nu \leq \frac{1}{1-\mu}} \mathrm{E}_F[\log(1 - (X - \mu)\nu)] \ , \tag{12}$$

where $\mathcal{V}(\tilde{\boldsymbol{M}})$ is the set of positive measures with $0, 1, \cdots, d$-th moments equal to $\tilde{\boldsymbol{M}}$, written in (16).

Proof (of Theorem 3). Let $\mathcal{M}_{d+1}$ be the moment space with respect to the system $(1, x, \cdots, x^d)$. It is easily checked that the solutions of (6) have one-to-one correspondence to the representations of $\tilde{\boldsymbol{M}} \in \mathcal{M}_{d+1}$ with index at most $d/2$ or to the upper principal representation.

First consider the trivial case that $\tilde{\boldsymbol{M}}$ is a boundary point of $\mathcal{M}_{d+1}$. For this case, the proof is straightforward since $\mathcal{V}(\tilde{\boldsymbol{M}})$ has a single element and its index is at most $d/2$ from Prop. 5.

Now we consider the remaining case that $\tilde{\boldsymbol{M}}$ is an interior point of $\mathcal{M}_{d+1}$. For this case, the upper principal representation of $\boldsymbol{M}$ is the unique solution of (6) since existence of a representation with index at most $d/2$ implies that $\tilde{\boldsymbol{M}}$ is a boundary point of $\mathcal{M}_{d+1}$ from Prop. 5. In the following, we complete the proof by showing that the upper principal representation of $\tilde{\boldsymbol{M}}$ is the optimal solution $\bar{F}$ in (5).

Consider applying the minimax theorem to (12). First, $\mathcal{F} \supset \mathcal{V}(\tilde{\boldsymbol{M}})$ is compact with respect to the Lévy distance and $\mathrm{E}_F[\log(1 - (X - \mu)\nu)]$ is linear in $F \in \mathcal{V}$ for any fixed ν. Next, $\mathrm{E}_F[\log(1 - (X - \mu)\nu)]$ is upper-semicontinuous and concave in ν for any fixed F. Then we obtain from the the minimax theorem that there exists $\bar{\nu}$ satisfying

$$D_{\min}(\boldsymbol{M}, \mu) = \inf_{F \in \mathcal{V}(\tilde{\boldsymbol{M}})} \mathrm{E}_F[\log(1 - (X - \mu)\bar{\nu})] \ .$$

Now we show $\bar{\nu} < (1-\mu)^{-1}$ by contradiction. Assume $\bar{\nu} = (1-\mu)^{-1}$. From Prop. 6 with $x^* := 1$, there exists $F' \in \mathcal{V}(\tilde{\boldsymbol{M}})$ such that $F'(\{1\}) > 0$. Therefore, from $\log 0 = -\infty$,

$$\inf_{F \in \mathcal{V}(\boldsymbol{M})} \mathrm{E}_F[\log(1 - (X-\mu)\bar{\nu})] \le \mathrm{E}_{F'}[\log(1-(X-\mu)\bar{\nu})] = -\infty \ .$$

It contradicts the positivity of $D_{\min}^{(d)}$ and $\bar{\nu} < (1-\mu)^{-1}$ is obtained.

From Lemma 3 with $p := 1+\mu\bar{\nu}$ and $q := \bar{\nu}$, $(1, x, \cdots, x^d, -\log(1-(x-\mu)\bar{\nu}))$ is a T-system on $[0,1]$ since $p/q = 1/\bar{\nu}+\mu > 1$. Therefore, from Prop. 8, we obtain

$$\begin{aligned}
\inf_{F \in \mathcal{V}(\boldsymbol{M})} \mathrm{E}_F[\log(1 - (X-\mu)\bar{\nu})] &= - \sup_{F \in \mathcal{V}(\boldsymbol{M})} \mathrm{E}_F[-\log(1-(X-\mu)\bar{\nu})] \\
&= -\mathrm{E}_{\bar{F}}[-\log(1-(X-\mu)\bar{\nu})] \\
&= \mathrm{E}_{\bar{F}}[\log(1-(X-\mu)\bar{\nu})] = D_{\min}(\bar{F}, \mu) \ ,
\end{aligned}$$

where $\bar{F}$ corresponds to the upper principal representation of $\tilde{\boldsymbol{M}}$. □

6 Improvement of DMED-M Policy

In DMED-M, $D_{\min}(F, \mu)$ is bounded from below by $D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu)$. When the gap between $D_{\min}^{(d)}$ and $D_{\min}$ is small, DMED-M behaves like the asymptotically optimal policy, DMED. In this section, we propose DMED-MM policy which is obtained by a slight modification to DMED-M. We discuss that DMED-MM works successfully for the case where the gap between $D_{\min}^{(d)}$ and $D_{\min}$ is large.

Define a function $\tilde{D}_{\min}^{(d)}(F, \mu)$ by

$$\tilde{D}_{\min}^{(d)}(F, \mu) = \begin{cases} D_{\min}(F, \mu) & \mathrm{E}_F\left[\frac{1}{1-X}\right] \le \frac{1}{1-\mu}, \\ D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu) & \text{otherwise,} \end{cases}$$

where recall that $D_{\min}(F, \mu) = \mathrm{E}_F[\log(1-X)] - \log(1-\mu)$ for the first case. *DMED-MM (DMED-M Mixed)* policy is obtained by replacing $D_{\min}^{(d)}(\mathrm{E}^{(d)}(\hat{F}_i(n)), \mu)$ in DMED-M by $\tilde{D}_{\min}^{(d)}(\hat{F}_i(n), \mu)$. Then the criterion for choosing an arm is the same as DMED for the case $\mathrm{E}_{\hat{F}_i(n)}[1/(1-X)] \le 1/(1-\mu)$ and the same as DMED-M otherwise.

In the first place, $D_{\min}^{(d)}(\mathrm{E}^{(d)}(\hat{F}_i(n)), \mu)$ in DMED-M is easy to compute since the empirical moments $\mathrm{E}^{(d)}(\hat{F}_i(n))$ can be obtained in constant time from the sum $\sum_{t=1}^{T_i(n)} X_{i,t}^m$ where $X_{i,t}$ denotes the t-th reward from arm i. On the other hand, $D_{\min}(\hat{F}_i(n), \mu) = \max_\nu \mathrm{E}_{\hat{F}_i(n)}[1-(X-\mu)\nu]$ in DMED requires the computation of the sum $\sum_t \log(1-(X_{i,t}-\mu)\nu)$ where μ and ν generally take various values. In this viewpoint, the computation of $\tilde{D}_{\min}^{(d)}(\hat{F}_i(n), \mu)$ is practical since it is obtained from the sums $\sum_t X_{i,t}^m$, $\sum_t 1/(1-X_{i,t})$ and $\sum_t \log(1-X_{i,t})$.

Now consider the maximum gap between $D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu)$ and $D_{\min}(F, \mu)$ among distributions F with moment $\boldsymbol{M}$.

Lemma 2. *The supremum in*

$$\sup_{F \in \mathcal{F}(\boldsymbol{M})} \left(D_{\min}(F, \mu) - D_{\min}^{(d)}(\boldsymbol{M}, \mu) \right) = \sup_{F \in \mathcal{F}(\boldsymbol{M})} D_{\min}(F, \mu) - D_{\min}^{(d)}(\boldsymbol{M}, \mu) \quad (13)$$

is attained by the unique solution of

$$\begin{cases} \sum_{i=1}^{\lceil (d+1)/2 \rceil} f_i x_i^m = M_m \; (m = 0, \cdots, d), & d \text{ is odd}, \\ \sum_{i=1}^{\lceil (d+1)/2 \rceil} f_i x_i^m = M_m \; (m = 0, \cdots, d), \; x_1 = 0, & d \text{ is even}. \end{cases} \quad (14)$$

The proof of the lemma is parallel to that of Theorem 3 which considers the infimum of $D_{\min}(F, \mu)$. In Theorem 3, we saw that the upper principal representation $\bar{F}$ of $\boldsymbol{M}$ attains the infimum from Prop. 8. Similarly, we can show that the lower principal representation $\underline{F}$ of $\boldsymbol{M}$ attains the supremum in (13).

Note that we can obtain an explicit expression of (13) for small d in the same way as Theorem 4. However, such an expression is a complicated function on $\boldsymbol{M}$ as Theorem 4 and it is not useful as an evaluation of the gap, since we cannot know the value of the expression until we substitute the specific value of $\boldsymbol{M}$.

Lemma 2 is useful when we consider the performance of DMED-MM. Compare the solution $\underline{F}$ of (14) to the upper principal representation $\bar{F}$ in (6). For odd d, $\underline{F}$ is supported by fewer points which generally contain neither 0 nor 1. For even d, $\underline{F}$ is supported by the same number of points which contain 0 instead of 1. In any case, we can say qualitatively that $\underline{F}$, distribution such that DMED-M behaves badly (i.e., gap between $D_{\min}$ and $D_{\min}^{(d)}$ is large), has small weight around 1.

Note that such a distribution often satisfies $\mathrm{E}_F[1/(1-X)] \le 1/(1-\mu)$ since $\mathrm{E}_F[1/(1-X)]$ is controlled mainly by the weight around 1. In fact, we can show from Prop. 8 that $\min_{F \in \mathcal{F}(\boldsymbol{M})} \mathrm{E}_F[1/(1-X)]$ is attained by the lower principal representation $\underline{F}$, which also attains the supremum in (13).

Now we summarize the above argument: (1) DMED-M behaves most differently from DMED for $\underline{F}$ among distributions with the moment $\boldsymbol{M}$. (2) Among these distributions F, $\underline{F}$ is also the distribution minimizing $\mathrm{E}_F[1/(1-X)]$, although the minimum value is not always smaller than $1/(1-\mu)$. (3) If $\mathrm{E}_F[1/(1-X)] \le 1/(1-\mu)$, DMED-MM behaves in the same way as DMED (otherwise it does in the same way as DMED-M). In this sense, the worst case gap between DMED-MM and DMED is sometimes smaller than that between DMED-M and DMED.

7 Experiments

In this section we show some numerical results on DMED-MM and the function $D_{\min}^{(d)}$.

First, we compare the performance of $1, 2, 3$-th degree DMED-MM and DMED in Fig. 1. Each plot is an average over 1000 different runs and we used 5 arms with beta distributions $\mathrm{Be}(\alpha, \beta)$. Note that beta distribution covers various forms of distributions on $[0, 1]$. The parameters of the arms are $(\alpha, \beta) = (9, 1), (0.7, 0.3),$
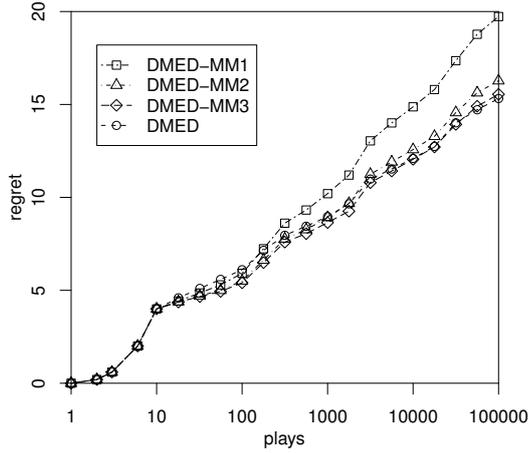
Fig. 1. Empirical regrets of $1, 2, 3$-th degree DMED-MM and DMED for beta distributions. Each plot is an average over 1000 different runs.

Table 1. Values of $D_{\min}^{(d)}(\mathrm{E}^{(d)}(F), \mu)$ and $D_{\min}(F, \mu)$ for beta distributions.

F	$\mathrm{E}_F[X]$	μ	$D_{\min}^{(1)}$	$D_{\min}^{(2)}$	$D_{\min}^{(3)}$	$D_{\min}$	$\mathrm{E}_F\left[\frac{1}{1-X}\right] \leq \frac{1}{1-\mu}$
$\mathrm{Be}(2,2)$	0.5	0.6	0.0204	0.0703	0.0843	0.0984	False
$\mathrm{Be}(0.5,0.5)$	0.5	0.6	0.0204	0.0366	0.0400	0.0408	False
$\mathrm{Be}(1,3)$	0.25	0.6	0.253	0.459	0.522	0.583	True
$\mathrm{Be}(0.25,0.75)$	0.25	0.6	0.253	0.348	0.391	0.431	False
$\mathrm{Be}(2,2)/2$	0.25	0.3	0.00617	0.0373	0.0490	0.0576	True
$\mathrm{Be}(0.5,0.5)/2$	0.25	0.3	0.00617	0.0239	0.0337	0.0401	True

$(5, 5), (0.3, 0.7), (1, 9)$ with expectations $\mu = 0.9, 0.7, 0.5, 0.3, 0.1$. The vertical axis denotes the regret $\sum_{i:\mu_i < \mu^*}(\mu^* - \mu_i)T_i(n)$, which is the loss due to choosing suboptimal arms. We see from the figure that the performance of DMED-MM approaches DMED as the degree increases.

Next, we show values of $D_{\min}^{(d)}$ and whether $\mathrm{E}_F[1/(1-X)] \leq 1/(1-\mu)$ or not for various distributions in Table 1. Recall that DMED-MM works the same as DMED for the case $\mathrm{E}_F[1/(1-X)] \leq 1/(1-\mu)$. Distribution $\mathrm{Be}(\alpha, \beta)/2$ denotes the distribution of $X/2$ for random variable X with distribution $\mathrm{Be}(\alpha, \beta)$. It corresponds to the case that the upper bound of the support of distributions is unknown and assumed conservatively as 2 instead of 1. For this case, a reward X is passed as $X/2$ to a policy for distributions on $[0, 1]$. We see from the figure that $D_{\min}^{(d)}$ bounds $D_{\min}$ from below accurately when $\mathrm{E}_F[1/(1-X)] \leq 1/(1-\mu)$

is false, as discussed in the previous section. Overall, the gap between $D_{\min}^{(1)}$ and $D_{\min}^{(2)}$ seems to be very large and it seems to be necessary to use at least the second moment (i.e., variance) to achieve a smaller regret.

8   Conclusion

In this paper we proposed DMED-M policy which is computed by the first d empirical moments of the arms. The regret bound of DMED-M approaches that of DMED, which is asymptotically optimal, as d increases. The computation involved in DMED-M is represented in an explicit form for small d. We also proposed DMED-MM policy, which sometimes improves the worst case performance of DMED-M.

An open problem is whether the asymptotic bound of DMED-M is the best for all policies which only consider the empirical moments. We may be able to prove the optimality of DMED-M in this sense under some regularity conditions.

References

[1] Lai, T.L., Robbins, H.: Asymptotically efficient adaptive allocation rules. Advances in Applied Mathematics **6** (1985) 4–22

[2] Agrawal, R.: The continuum-armed bandit problem. SIAM J. Control Optim. **33** (1995) 1926–1951

[3] Even-Dar, E., Mannor, S., Mansour, Y.: Pac bounds for multi-armed bandit and markov decision processes. In: Proceedings of COLT 2002, London, UK, Springer-Verlag (2002) 255–270

[4] Meuleau, N., Bourgine, P.: Exploration of multi-state environments: Local measures and back-propagation of uncertainty. Machine Learning **35** (1999) 117–154

[5] Strens, M.: A bayesian framework for reinforcement learning. In: Proceedings of ICML 2000, Morgan Kaufmann, San Francisco, CA (2000) 943–950

[6] Vermorel, J., Mohri, M.: Multi-armed bandit algorithms and empirical evaluation. In: Proceedings of ECML 2005, Porto, Portugal, Springer (2005) 437–448

[7] Gittins, J.C.: Multi-armed bandit allocation indices. Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons Ltd., Chichester (1989)

[8] Auer, P., Cesa-Bianchi, N., Freund, Y., Schapire, R.E.: The nonstochastic multi-armed bandit problem. SIAM J. Comput. **32** (2002) 48–77

[9] Uchiya, T., Nakamura, A., Kudo, M.: Algorithms for adversarial bandit problems with multiple plays. In: Proceedings of ALT 2010, Berlin, Heidelberg, Springer-Verlag (2010) 375–389

[10] Auer, P., Cesa-Bianchi, N., Fischer, P.: Finite-time analysis of the multiarmed bandit problem. Machine Learning **47** (2002) 235–256

[11] Audibert, J.Y., Munos, R., Szepesvári, C.: Exploration-exploitation tradeoff using variance estimates in multi-armed bandits. Theor. Comput. Sci. **410** (2009) 1876–1902

[12] Honda, J., Takemura, A.: An asymtotically optimal bandit algotirhm for bounded support models. In: Proceedings of COLT 2010, Haifa, Israel (2010) 67–79

[13] Burnetas, A.N., Katehakis, M.N.: Optimal adaptive policies for sequential allocation problems. Adv. Appl. Math. **17** (1996) 122–142

[14] Lamperti, J.: Probability; a survey of the mathematical theory. Second edn. Wiley Series in Probability Statistics. John Wiley & Sons Ltd., New York (1996)

[15] Neumann, J.V.: Zur Theorie der Gesellschaftsspiele. Mathematische Annalen **100** (1928) 295–320

[16] Karlin, S., Studden, William J, .: Tchebycheff systems, with applications in analysis and statistics. Interscience Publishers New York, (1966)

A Tchebycheff Systems and Moment Spaces

In this appendix we summarize results on Tchebycheff systems and moment spaces needed for the proof of Theorem 3. All functions and measures are defined on $[a, b]$ $(a < b)$ in this appendix whereas they are on $[0, 1]$ elsewhere. For any set of points $\{x_1, \cdots, x_l\}$, we always assume $a \le x_1 < x_2 < \cdots < x_l \le b$.

Definition 1. *Let* $u_0(x), \cdots, u_d(x)$ *denote continuous real-valued functions on* $[a, b]$. *These functions are called a Tchebycheff system (or T-system) if determinants*

$$
\det \begin{pmatrix} u_0(x_0) \; u_1(x_0) \; \cdots \; u_d(x_d) \\ u_0(x_1) \; u_1(x_1) \; \cdots \; u_d(x_d) \\ \vdots \qquad \vdots \qquad\qquad \vdots \\ u_0(x_d) \; u_1(x_d) \; \cdots \; u_d(x_d) \end{pmatrix}
\tag{15}
$$

are positive for all $\{x_0, \cdots, x_d\}$.

A typical T-system is $u_i(x) = x^i$ $(i = 0, 1, \cdots, d)$, where (15) is represented as the Vandermonde determinant

$$
\det \begin{pmatrix} 1 \; x_0 \; \cdots \; x_0^d \\ 1 \; x_1 \; \cdots \; x_1^d \\ \vdots \; \vdots \qquad \vdots \\ 1 \; x_d \; \cdots \; x_d^d \end{pmatrix} = \prod_{1 \le i < j \le d} (x_j - x_i) > 0 \ .
$$

Let $Z(u)$ of a function $u(x)$ denote the number of distinct points $x \in [a, b]$ such that $u(x) = 0$. Then T-systems are discriminated by the following proposition.

Proposition 4 ([16, Chap. I, Theorem 4.1]). *If a system* $\{u_i\}_{i=0}^d$ *of continuous functions on* $[a, b]$ *satisfies* $Z(u) \le d$ *for all*

$$
u(x) = \sum_{i=0}^d a_i u_i(x), \ \{a_i\} \in \mathbb{R}^{d+1} \setminus \{0^{d+1}\} \ ,
$$

then $(u_0, u_1, \cdots, u_{d-1}, u_d)$ *or* $(u_0, u_1, \cdots, u_{d-1}, -u_d)$ *is a T-system.*

Lemma 3. *For any p and* $q > 0$ *satisfying* $b < p/q$, $(1, x, \cdots, x^d, -\log(p - qx))$ *is a T-system on* $[a, b]$.

Proof. Let $b' \in (b, p/q)$ be sufficiently close to p/q and consider function

$$u(x) = \sum_{m=0}^{d} a_m x^m - a_{d+1} \log(p - qx)$$

on $x \in [a, b']$. Since the derivative of $u(x)$ is written as

$$\frac{\mathrm{d}u(x)}{\mathrm{d}x} = \frac{(p - qx) \sum_{m=1}^{d} a_m x^{m-1} + a_{d+1}q}{p - qx} ,$$

$u(x)$ has at most d extreme points in $[a, b']$. Therefore $Z(u) \leq d + 1$ and $(1, x, \cdots, x^d, \log(p - qx))$ or $(1, x, \cdots, x^d, -\log(p - qx))$ is a T-system on $[a, b']$ from Prop. 4.

The determinant (15) for the system $(1, x, \cdots, x^d, \log(p - qx))$ is written as

$$\det \begin{pmatrix} 1 & x_0 & \cdots & x_0^d & \log(p - qx_0) \\ 1 & x_1 & \cdots & x_1^d & \log(p - qx_1) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & x_{d+1} & \cdots & x_{d+1}^d & \log(p - qx_{d+1}) \end{pmatrix}$$

$$= \sum_{m=0}^{d+1} (-1)^{d+m+1} \left(\prod_{0 \leq i < j \leq d+1 : i, j \neq m} (x_j - x_i) \right) \log(p - qx_m) .$$

For the case that $x_{d+1} = b'$ with $b' \uparrow p/q$, $\log(p - qx_{d+1})$ goes to $-\infty$ and the sign of the determinant is controlled by the term involving $\log(p - qx_{d+1})$, which is written as

$$(-1)^{2m+2} \left(\prod_{0 \leq i < j \leq d+1 : i, j \neq m} (x_j - x_i) \right) \log(p - qx_{d+1}) < 0 .$$

Then, $(1, x, \cdots, x^d, \log(p - qx))$ cannot be a T-system on $[a, b']$ for b' sufficiently close to p/q and therefore $(1, x, \cdots, x^d, -\log(p - qx))$ has to be a T-system on $[a, b']$. From the definition of T-system, it also is a T-system on $[a, b] \subset [a, b']$. $\square$

Let $\mathcal{V}$ be the family of positive measures on $[a, b]$ and define a subset $\mathcal{V}(\tilde{\boldsymbol{M}})$ of $\mathcal{V}$ for a vector $\tilde{\boldsymbol{M}} = (M_0, M_1, \cdots, M_d)$ as

$$\mathcal{V}(\tilde{\boldsymbol{M}}) = \left\{ \sigma \in \mathcal{V} : \forall m \in \{0, 1, \cdots, d\}, \int_a^b x^m \mathrm{d}\sigma(x) = M_m \right\} . \tag{16}$$

The notion of *moment spaces* is essential to examine properties of T-systems.

Definition 2. *The moment space $\mathcal{M}_{d+1}$ with respect to the T-system $\{u_i\}$ is given by*

$$\mathcal{M}_{d+1}$$
$$\equiv \left\{ \left(\int_a^b u_0(x) \mathrm{d}\sigma(x), \int_a^b u_1(x) \mathrm{d}\sigma(x), \cdots, \int_a^b u_d(x) \mathrm{d}\sigma(x) \right) \in \mathbb{R}^{d+1} : \sigma \in \mathcal{V} \right\} .$$

Consider the case that $\tilde{\boldsymbol{M}} \in \mathcal{M}_{d+1}$ satisfies

$$M_m = \sum_{i=1}^{l} f_i u_m(x_i) \quad (m = 0, \cdots, d) \tag{17}$$

with $x_1, \cdots, x_l \in [a, b]$ and $f_1, \cdots, f_l > 0$ for any finite l. We call such an expression *representation* of $\tilde{\boldsymbol{M}}$. A representation of $\tilde{\boldsymbol{M}}$ corresponds uniquely to the measure

$$\sigma = \sum_{i=1}^{l} f_i \delta(x_i) \in \mathcal{V}$$

for the delta measure $\delta(x)$ at point x. We sometimes identify the measure σ with the representation of $\tilde{\boldsymbol{M}}$. The measure σ is a probability measure if $\sum_i f_i = 1$.

The *index* of the representation (17) is defined as the number of the points $(x_1, \cdots, x_l)$ under the special convention that the points a, b are counted as one half. A representation is called *principal* if its index is $(d+1)/2$. Furthermore, the representation is *upper* if $(x_1, \cdots, x_l)$ contains b and *lower* otherwise.

For the proof of Theorem 3, it is necessary to study the nature on the set $\mathcal{V}(\tilde{\boldsymbol{M}})$. It differs according to whether $\tilde{\boldsymbol{M}}$ is a boundary point of $\mathcal{M}_{d+1}$ or an interior point of $\mathcal{M}_{d+1}$.

Proposition 5 ([16, Chap. II, Theorem 2.1]). *Every boundary point $\tilde{\boldsymbol{M}}$ has a unique representation. Moreover, $\tilde{\boldsymbol{M}} \in \mathcal{M}_{d+1}$ is a boundary point of $\mathcal{M}_{d+1}$ if and only if there exists a representation of $\tilde{\boldsymbol{M}}$ with index at most $d/2$.*

Proposition 6 ([16, Chap. II, Theorem 3.1]). *If $\tilde{\boldsymbol{M}}$ is an interior point of $\mathcal{M}_{d+1}$ then, for arbitrary $x^* \in [a, b]$, there exists a representation of $\tilde{\boldsymbol{M}}$ such that $(x_1, \cdots, x_l)$ contains x^*.*

Proposition 7 ([16, Chap. II, Corollary 3.1]). *If $\tilde{\boldsymbol{M}}$ is an interior point of $\mathcal{M}_{d+1}$ then there exist precisely one upper and one lower principal representations of $\tilde{\boldsymbol{M}}$.*

We use Prop. 7 implicitly in Prop. 8 below and the proof of Theorem 3. Prop. 8 is the main result of the appendix.

Proposition 8 ([16, Chap. III, Theorem 1.1]). *Assume $(u_0, u_1, \cdots, u_d)$ and $(u_0, u_1, \cdots, u_d, h)$ are T-systems. Then*

$$\max_{\sigma \in \mathcal{V}(\tilde{\boldsymbol{M}})} \int_a^b h(x) \mathrm{d}\sigma(x)$$

is attained uniquely by $\bar{\sigma}$, the upper principal representation of $\tilde{\boldsymbol{M}}$. Similarly,

$$\min_{\sigma \in \mathcal{V}(\tilde{\boldsymbol{M}})} \int_a^b h(x) \mathrm{d}\sigma(x)$$

is attained uniquely by $\underline{\sigma}$, the lower principal representation of $\tilde{\boldsymbol{M}}$.