We begin by recalling properties of the measures \( E \) and \( Q \) introduced in Section 3. By Proposition 7.28 in Bertsekas and Shreve [1996], the measures \( P_t^{A,\mu} \), \( t \in \mathbb{Z}_+ \), and \( P_t^{A,\mu} \) satisfy the following properties (see also Proposition V.1.1 of [Neveu 1965]):

1. For every real-valued function \( q \) that is integrable on \((\Omega_t, F_t, P_t^{A,\mu})\), we have
   \[
   \int_{\Omega_t} q(h_t) P_t^{A,\mu}(dh_t) = \int \int q(u, s, y) \times Q^{\mu}(dy|s) \pi_1(ds|u) \lambda(du),
   \]
   (12)

2. For every \( t > 1 \) and every integrable real-valued function \( q \) on \((\Omega_t, F_t, P_t^{A,\mu})\), we have
   \[
   \int_{\Omega_t} q(h_t) P_t^{A,\mu}(dh_t) = \int_{\Omega_{t-1}} \int_{[0,1]^n} \int_{\mathcal{D}} q(h_{t-1}, u, s, y) \times Q^{\mu}(dy|s) \pi_1(ds|u) \lambda(du) P_{t-1}^{A,\mu}(dh_{t-1}).
   \]
   (13)

3. For every \( t \in \mathbb{Z}_+ \) and every Borel subset \( A \) of \( \Omega_t \), we have
   \[
   P_t^{A,\mu}(A) = P_t^{A,\mu}(A \times S \times S \times \cdots).
   \]

Next, suppose Assumption 2 holds, and let \( \rho_{\varepsilon,\nu}(\cdot) \) denote the Gaussian measure on \( \mathbb{R} \) with mean \( \varepsilon \in \mathbb{R} \) and standard deviation \( \nu > 0 \). Then, for every \( \zeta \in \mathbb{R}^I \) and \( s \in \mathcal{D} \), the measure \( Q^\zeta(\cdot|s) \) is a Gaussian measure on \( \mathbb{R} \) having density \( \rho_{\varepsilon(s),\sigma}(\cdot) \) w.r.t. the Lebesgue measure on \( \mathbb{R} \). Consequently, for every \( \zeta \in \mathbb{R}^I \) and \( s \in \mathcal{D} \), the measures \( Q^\zeta(\cdot|s) \) are mutually absolutely continuous, and
   \[
   \frac{dQ^\zeta(\cdot|s)}{dQ^\nu(\cdot|s)} = \frac{\rho_{\varepsilon(s),\sigma}(y)}{\rho_{\varepsilon(s),\sigma}(y)}. \tag{14}
   \]

We are now ready to begin the proof of Theorem 3.2.

**Proof of Theorem 3.2** Consider an alternative reward model given by \( \zeta \in Alt_\varepsilon(\mu) \). For each \( t \in \mathbb{Z}_+ \), define the log-likelihood ratio \( L_t^{\mu,\zeta} \) \( \defeq \ln \frac{dP_t^{A,\mu}}{dP_t^{A,\nu}} \), and note that \( L_t^{\mu,\zeta} \) is a random variable on \((\Omega_t, F_t, P_t^{A,\mu})\). It is now easy to see from (12), (13) and (14) that, for each \( t \in \mathbb{Z}_+ \) and each \( h_t = \{(s_i, y_i, u_i)\}_{i=1}^t \in \Omega_t \), we have
   \[
   L_t^{\mu,\zeta}(h_t) = \sum_{i=1}^t \ln \frac{\rho_{g_{s_i}(s_i),\sigma}(y_i)}{\rho_{g_{s_i}(s_i),\sigma}(y_i)}. \tag{15}
   \]

Next, define the event \( \mathcal{E}^{\mu} \) \( \defeq \{ \arg \max_{s \in \mathcal{D}} g_{h_{\tau}}(s) \subseteq \mathcal{F}(h_{\tau}) \subseteq O_\varepsilon(\mu) \} \), and note that \( \mathcal{E}^{\mu} \) is contained in the \( \sigma \)-algebra \( \mathcal{F}_{\tau} \) generated by the stopping time \( \tau \). It follows from Lemma 19 of Kaufmann et al. [2016] that
   \[
   E^{A,\mu}(L_t^{\mu,\zeta}) \geq \text{kl}(P_t^{A,\mu}(\mathcal{E}^{\mu}), P_t^{A,\zeta}(\mathcal{E}^{\mu})), \tag{16}
   \]
   where \( \text{kl}(\nu_1, \nu_2) \) is the KL-divergence between two Bernoulli distributions having parameters \( \nu_1, \nu_2 \in [0, 1] \).

Next, define the event \( \mathcal{E}^{\zeta} \) by replacing \( \mu \) in the definition of \( \mathcal{E}^{\mu} \) with \( \zeta \). Since \( A \) is a \((\varepsilon, \delta)\)-PAC algorithm, we have \( \text{kl}(P_t^{A,\mu}(\mathcal{E}^{\mu}), P_t^{A,\zeta}(\mathcal{E}^{\mu})) \geq 1 - \delta \) and \( \text{kl}(P_t^{A,\zeta}(\mathcal{E}^{\zeta}), \mathcal{O}_\varepsilon(\mathcal{E}^{\zeta})) = 0 \). As a result, we infer that \( \text{kl}(P_t^{A,\zeta}(\mathcal{E}^{\mu}), P_t^{A,\zeta}(\mathcal{E}^{\mu}))) \leq \delta \). By choice of \( \zeta \), we have \( \mathcal{E}^{\mu} \cap \mathcal{E}^{\zeta} \subseteq \{ \mathcal{F}(h_{\tau}) \subseteq O_\varepsilon(\mu) \cap O_\varepsilon(\zeta) \} = \emptyset \). \( \text{kl}(\delta, 1 - \delta) \) by inequality (3) in Kaufmann et al. [2016], we further have \( \text{kl}(\delta, 1 - \delta) \geq \ln(1/2.48) \). Using this in (16), we get
   \[
   E^{A,\mu}(L_t^{\mu,\zeta}) \geq \frac{1}{2.48}. \tag{17}
   \]
Combining (17) with (18) from Lemma A.1 below yields
\[
\frac{1}{2\sigma^2} \mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2 \geq \ln \left( \frac{1}{2\delta_0} \right).
\]

Inequality (2) now follows by taking an infimum over $\zeta \in A_{1,\zeta}(\mu)$ on the left hand side in the inequality above and rearranging the resulting inequality. □

**Lemma A.1.** Let algorithm $A$ and $\mu, \zeta \in \mathbb{R}^f$ be as in the proof of Theorem 4.2. Suppose \( \{L_t^{\mu, \zeta}\}_{t=1}^\infty \) is defined as in (15), and let $\tau$ be a stopping time with respect to the filtration $\{F_t\}_{t=0}^\infty$. Then we have
\[
\mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2 \leq \frac{1}{2\sigma^2} \mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2.
\]

**Proof.** For each $t \in \mathbb{Z}_+$, define $\ell_t = \ln \rho_{g_\mu, (y_t)} - \ln \rho_{g_\zeta, (y_t)}$ and let $G_t$ denote the $\sigma$-algebra on $\Omega$ generated by $\{h_{t-1, u_t, s_t}\}$. Note that, for each $t \in \mathbb{Z}_+$, $\ell_t$ is a $F_t$-measurable random variable, while $F_{t-1} \subseteq G_t \subseteq F_t$.

Next, define the process \( \{M_t\}_{t=0}^\infty \) by $M_0 = 0$ and $M_t = \sum_{i=1}^t (\ell_i - \mathbb{E}^A(\ell_i|G_i))$ for each $t \in \mathbb{Z}_+$. The inclusions $F_{t-1} \subseteq G_t \subseteq F_t$ along with the tower property of conditional expectations show that the process \( \{M_t\}_{t=0}^\infty \) is adapted to the filtration $\{F_t\}_{t=0}^\infty$ and is a martingale under the measure $\mathbb{P}^{A, \mu}$. The optional stopping theorem now implies that $\mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2 \leq \mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2$. This immediately yields
\[
\mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2 \leq \mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2.
\]

Substituting the expression for a Gaussian density in the expression for $\ell_t$ yields $2\sigma^2\ell_t = 2\|g_\mu(s_t) - g_\zeta(s_t)\|^2 - \|g_\zeta(s_t)^2\|^2$ for each $i \in \mathbb{Z}_+$. Using the fact that $\mathbb{E}^A(\ell_i) = \mathbb{E}^A(\ell_i|G_i) = \|g_\zeta(s_t) - g_\mu(s_t)\|^2 \leq \mathbb{E}^A(\mathbb{E}^A(\ell_i|G_i))$ for each $i \in \mathbb{Z}_+$. Using the last inequality in (19) and recognizing the left hand side of (19) to be $\mathbb{E}^A(\tau)$ (8), we have
\[
\mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2 \leq \mathbb{E}^A(\tau)\|g_\mu - g_\zeta\|_\infty^2.
\]

**C PROOF OF PROPOSITION 4.3**

By way of preparation for the proof of Proposition 4.3, we will find it convenient to rewrite (3) and (4) by grouping together observations made during each round. To this end, let $B_{L,m} = [\phi(p_1), \ldots, \phi(p_m)] \in \mathbb{R}^{f \times m}$ and, for each $i \in \mathbb{Z}_+$, let $y_i^j = [y_{(i-1),m+1}, \ldots, y_{jm}]^T \in \mathbb{R}^m$ and $\eta_i^j = [\eta_{(i-1),m+1}, \ldots, \eta_{jm}]^T \in \mathbb{R}^m$. Denote the vectors of rewards and noise samples, respectively, encountered in the $j$th round. The decision epoch at the end of $k > 0$ rounds is $t = km$. In the notation of (3), we have
\[
X_t = [B_{L,m}]_{k \text{ times}} = [B_{L,m}]_{k \text{ times}}.
\]

Equations (3)-(4) now become
\[
\mu_{km} = (B_{L,m}B_{L,m}^T)^{-1}B_{L,m}\left[\frac{1}{k}\sum_{j=1}^k y_i^j\right],
\]
\[
\mu_{km} - \mu = (B_{L,m}B_{L,m}^T)^{-1}B_{L,m}\left[\frac{1}{k}\sum_{j=1}^k \eta_i^j\right].
\]

The proof of the sub-Gaussian part of Proposition 4.3 essentially applies to the right hand side of (21), the tail concentration inequality below for the norm of the average of $k$ random vectors having independent $\sigma$-sub-Gaussian components. The proof is given later in this appendix.
Proposition C.1. Suppose $\xi_1, \ldots, \xi^k$ are $f$-dimensional random vectors such that the random variables $\{\xi_i^j : i = 1, \ldots, f, j = 1, \ldots, k\}$ are independent and $\sigma$-sub Gaussian. Let $S_k = (\xi^1 + \cdots + \xi^k)$. Then the following statements hold.

1. $\exp (\lambda \|S_k\|^2)$ is integrable for each $\lambda \in (0, 1/2\sigma^2 k)$.
2. For every $\epsilon > 0$, we have $P \left( \frac{1}{k} \|S_k\|_2 > \epsilon \right) \leq \beta(k, \epsilon)$, where $\beta$ is given by [2].

We are now ready to prove Proposition C.1.

Proof of Proposition C.1. First, suppose Assumption 1 holds. We have

$$
\|g_{\mu_{km}} - g_{\mu_m}\|_\infty = \max_{s \in D} |\phi^T(s)(\mu_{km} - \mu)| = \max_{s \in D} |\phi^T(s)B_{L,m}^T(B_{L,m}^T)^{-1}B_{L,m}(\frac{1}{k}S_k)|,
$$

where $S_k = \sum_{j=1}^k \bar{\eta}^j$, and the last equality uses (21). Since the columns of $B_{L,m}$ form a $(L, m)$-volumetric spanner for $\phi(D)$, it follows that $\|B_{L,m}^T(B_{L,m}^T)^{-1}\|_2 \leq L$ for all $s \in D$. Using this fact along with the Cauchy-Schwarz inequality in (22) gives $\|g_{\mu_{km}} - g_{\mu_m}\|_\infty \leq \frac{1}{\sqrt{k}} \|S_k\|_2$. The assertion of the proposition now follows immediately from Proposition C.1.

For proving Proposition C.1, we first recollect a few preliminary results. Though these results are known, we state them to make the constants explicit, and provide proofs for easy reference.

Lemma C.2. Suppose $X$ is $\sigma$-sub Gaussian for some $\sigma > 0$. If $\lambda \in (0, 1/\sigma)$, then $\exp (\lambda X^2)$ is integrable, and $E[\exp (\lambda X^2)] \leq 2^{2\sigma^2} (1 - \sigma^2)^{-1}$.

Proof. Let $\lambda \in (0, 1/\sigma)$. Since $X$ is $\sigma$-sub Gaussian, we have $P(|X| > t) \leq 2 \exp \left( -\frac{t^2}{2\sigma^2} \right)$ for all $t > 0$. Next, note that $\exp (\lambda X^2) \geq 1$. Let $s \geq 1$. Then

$$
P \left( \exp (\lambda X^2) > s \right) = P \left( |X| > \sqrt{\ln s} / \lambda \right) \leq 2 \exp \left( -\frac{1}{2\sigma^2} \ln s / \lambda \right) = 2s^{-\frac{1}{2\sigma^2}}.
$$

Thus, we conclude that

$$
P \left( \exp (\lambda X^2) > s \right) \leq \begin{cases} 1, & \text{if } s \leq 2^{2\sigma^2}\lambda, \\ 2s^{-\frac{1}{2\sigma^2}}, & \text{if } s > 2^{2\sigma^2}\lambda. \end{cases}
$$

Since $2\sigma^2 \lambda < 1$, the integral $\int_0^\infty P \left( \exp (\lambda x^2) > s \right) ds$ exists. Indeed, (23) implies that

$$
\int_0^\infty P \left( \exp (\lambda x^2) > s \right) ds 
\leq \int_0^{2^{2\sigma^2}\lambda} 1 ds + \int_{2^{2\sigma^2}\lambda}^\infty 2s^{-\frac{1}{2\sigma^2}} ds 
= \frac{2\sigma^2\lambda (1 - 2\sigma^2\lambda)^{-1}}{2\sigma^2}.
$$

Since $\exp (\lambda x^2)$ is a non-negative random variable, it follows that $E \left[ \exp (\lambda x^2) \right] = \int_0^\infty P \left( \exp (\lambda x^2) > s \right) ds$, and the result follows.

Lemma C.3. Suppose $\zeta$ is a random vector of dimension $f$ such that $\zeta_1, \ldots, \zeta_f$ are independent $\sigma$-sub Gaussian random variables. Then $E \left[ \exp (\lambda \zeta^T \zeta) \right] \leq \exp \left( \frac{\lambda^2 \|\zeta\|^2_2 \sigma^2}{2} \right)$ for all $\lambda \in \mathbb{R}$ and $\lambda \in \mathbb{R}$. Furthermore, if $\lambda \in (0, 1/2\sigma^2)$, then $\exp (\lambda \|\zeta\|^2_2)$ is integrable, and $E \left[ \exp (\lambda \|\zeta\|^2_2) \right] \leq \left( \frac{2^{2\sigma^2} - \lambda^2}{2} \right)^f$.

Proof. By independence and $\sigma$-sub Gaussianity, we have

$$
E \left[ \exp (\lambda X^T \zeta) \right] = \prod_{i=1}^f E \left[ \exp (\lambda X_i \zeta_i) \right] 
\leq \prod_{i=1}^f \exp \left( \frac{\lambda^2 \|X_i\|^2_2 \sigma^2}{2} \right) = \exp \left( \frac{\lambda^2 \|X\|^2_2 \sigma^2}{2} \right).
$$

This proves the first assertion. To prove the second assertion, let $\lambda \in (0, 1/2\sigma^2)$. By Lemma C.2 $\exp (\lambda \zeta_i^2)$ is integrable for each $i$. Hence it follows by independence that $\exp (\lambda \|\zeta\|^2_2)$ is also integrable, and $E \left[ \exp (\lambda \|\zeta\|^2_2) \right] = \prod_{i=1}^f E \left[ \exp (\lambda \zeta_i^2) \right] \leq \left( \frac{2^{2\sigma^2} - \lambda^2}{2} \right)^f$.

The next lemma, which we state without proof, is a conditional version of the first part of Lemma C.3.

Lemma C.4. Suppose $\zeta$ is a random vector of dimension $f$ such that $\zeta_1, \ldots, \zeta_f$ are independent, $\sigma$-sub Gaussian random variables. Let $Y$ be a $G$-measurable $f$-dimensional random vector, where $G$ is a $\sigma$-algebra such that $\zeta$ is independent of $G$. Then

$$
E \left[ \exp (\lambda Y^T \zeta) | G \right] \leq \exp \left( \frac{\lambda^2 \|Y\|^2_2 \sigma^2}{2} \right) a.s.
$$

The proof of Proposition C.1 follows next.

Proof of Proposition C.1. The $i$th component of $S_k$ is a sum of $k$ independent $\sigma$-sub Gaussian random variables. Applying the first part of Lemma C.3 with $\zeta = [\xi_1^i, \ldots, \xi_k^i]$ and $x = [1, \ldots, 1]$ lets us conclude that the $i$th element of $S_k$ is $\sigma \sqrt{k}$-sub Gaussian. Applying the second part of
Lemma C.3 with \( \zeta = S_k \) shows that \( \exp (\lambda \| S_k \|_2^2) \) is integrable for \( \lambda \in (0, 1/2\sigma^2 k) \). This proves the first assertion.

To prove the second statement, choose \( x \in \mathbb{R}^j \), and define the process \( \{ M_j(x) \}_{j=0}^k \) by \( M_0(x) = 1 \)

\[
M_j(x) = \exp \left( x^T S_j - \frac{\sigma^2}{2} \| x \|_2^2 \right), \quad j = 1, \ldots, k,
\]

where \( S_j = \xi^1 + \cdots + \xi^j \) for each \( j \). It follows from the first part of Lemma C.3 that \( M_j(x) \) is integrable for each \( j \). Next, let \( G_j \) denote the \( \sigma \)-algebra generated by \( \xi^1, \ldots, \xi^j \), with \( G_0 \) denoting the trivial \( \sigma \)-algebra, and note that \( M_j(x) \) is \( F_j \)-measurable. For each \( j = 1, \ldots, k \), we have

\[
\mathbb{E} [ M_j(x) | G_{j-1} ] = \mathbb{E} \left[ M_{j-1}(x) \exp \left( x^T \xi^j - \frac{\sigma^2}{2} \| x \|_2^2 \right) | G_{j-1} \right] = M_{j-1}(x) \mathbb{E} \left[ \exp \left( x^T \xi^j - \frac{\sigma^2}{2} \| x \|_2^2 \right) \right] \leq M_{j-1}(x),
\]

where the second equality follows from the \( G_{j-1} \)-measurability of \( M_{j-1}(x) \) and the \( G_{j-1} \)-independence of \( \xi^j \) (see Lemma C.4), while the last inequality follows by applying the first part of Lemma C.3 with \( \zeta = \xi^j \). Thus, \( \{ M_j(x) \}_{j=0}^k \) is a supermartingale with respect to the filtration \( \{ G_j \}_{j=0}^k \).

Next, define \( \{ M_j \}_{j=0}^k \) by

\[
M_j = \left( \frac{k \sigma^2}{2\pi} \right)^{j/2} \int_{\mathbb{R}^j} M_j(x) \exp \left( -\frac{k \sigma^2}{2} x^T x \right) dx,
\]

and note that \( M_0 = 1 \). Substituting for \( M_j(x) \) in (24), completing the square in the exponent and rearranging terms yields

\[
M_j = \left( \frac{k \sigma^2}{(j+k)\sigma^2} \right)^{j/2} \exp \left( \frac{\| S_j \|_2^2}{2(j+k)\sigma^2} \right) \times J.
\]

where \( J \) is the integral over \( x \) of the \( F \)-dimensional Gaussian density over \( x \) with mean \( [(j+k)\sigma^2]^{-1} S_j \) and covariance matrix \( [(j+k)\sigma^2]^{-1} I \), with \( I \) denoting the \( f \times f \) identity matrix. Thus, \( J \) evaluates to 1. Next, \( S_j \) is a random vector with independent \( \sigma \)-sub-Gaussian components. Also, \( \frac{1}{(j+k)\sigma^2} < \frac{1}{2\sigma^2} \). Hence, by Lemma C.3 \( M_j \) is integrable. In addition, it follows from Lemma 20.3 in Lattimore and Szepesvári (2020) that \( \{ M_j \}_{j=0}^k \) is a submartingale.

Letting \( j = k \) in (25) gives \( M_k = 2^{-k/2} \exp(\frac{\| S_k \|_2^2}{2\sigma^2}) \). By Ville’s maximal inequality (see Theorem 3.9 in Lattimore and Szepesvári (2020)), we have

\[
\mathbb{P}(\| S_k \|_2 > \epsilon) = \mathbb{P} \left( M_k > \frac{1}{2^{(k/2)}} \exp \left( \frac{\epsilon^2}{4k\sigma^2} \right) \right) \leq \mathbb{E} \left[ \max_j M_j > \frac{1}{2^{(k/2)}} \exp \left( \frac{\epsilon^2}{4k\sigma^2} \right) \right] \leq \frac{1}{2^{(k/2)}} \exp \left( \frac{-\epsilon^2}{4k\sigma^2} \right).
\]

Replacing \( \epsilon \) by \( k \epsilon \) in the last inequality completes the proof of the second assertion. \( \square \)

### D PROOF OF PROPOSITION 5.1

The proof of Proposition 5.1 uses the following lemma.

**Lemma D.1.** Let \( s \in [p_{\min}, p_{\max}] \) and suppose \( p_1, \ldots, p_f \in [p_{\min}, p_{\max}] \) are such that \( p_i \neq p_j \) for all \( i \neq j \). Then \( c_1, \ldots, c_{n+1} \in \mathbb{R} \) satisfy

\[
c_1 \phi(p_1) + \cdots + c_{n+1} \phi(p_f) = \phi(s) \tag{26}
\]

if and only if \( c_i = l_i(s, p) \) for each \( i = 1, \ldots, f \) where \( p = [p_1, \ldots, p_f]^T \), and \( l_i(s, p) \) is the \( i \)th Lagrange basis polynomial for the points \( \{ p_1, p_2, \ldots, p_f \} \) given by

\[
l_i(s, p) \overset{\text{def}}{=} \prod_{j \neq i} \frac{(s - p_j)}{(p_i - p_j)}. \tag{27}
\]

**Proof.** Equation (26) may be rewritten as

\[
V(p)c(s) = \phi(s), \tag{28}
\]

where \( V(p) \overset{\text{def}}{=} [\phi(p_1), \ldots, \phi(p_f)] \in \mathbb{R}^{f \times f} \). Note that \( V(p) \) is a Vandermonde matrix, and its determinant is given by (see Fact 7.18.5 from Bernstein (2018))

\[
det(V(p)) = \prod_{1 \leq i < j \leq f} (p_j - p_i), \tag{29}
\]

The determinant of \( V(p) \) in (29) is nonzero since \( p_i \neq p_j \) for \( j \neq i \). Equation (28) thus has a unique solution. Applying Cramer’s rule (see Fact 3.16.12 from Bernstein (2018)) gives this solution to be

\[
c_i = \frac{\text{det}(V(p^i))}{\text{det}(V(p))} \tag{30}
\]

where \( p^i \) is the vector obtained by replacing the \( i \)th element of \( p \) by \( s \). Using (29) to expand the determinants of the two Vandermonde matrices in (30) and canceling common terms gives \( c_i = l_i(s, p) \). \( \square \)
The proof of Proposition 5.1 follows.

**Proof of Proposition 5.1.** To show 1) implies 2), suppose \( p_1, \ldots, p_f \in \mathcal{D} \) are \((1,f)\)-volumetric points for the pair \((\phi, \mathcal{D})\). Choose \( s \in \mathcal{D} = [p_{\min}, p_{\max}] \) arbitrarily. Applying the definition of \((1,f)\)-volumetric points, it follows that there exist \( c_1, \ldots, c_f \in \mathbb{R} \) such that \( c_1 \phi(p_1) + \cdots + c_f \phi(p_f) = \phi(s) \) and \( c_1^2 + \cdots + c_f^2 \leq 1 \). Clearly, \( |c_i| \leq 1 \) for all \( i = 1, \ldots, f \). Since \( s \in \mathcal{D} \) was chosen arbitrarily, it follows that \( \{\phi(p_1), \ldots, \phi(p_f)\} \) is a barycentric spanner for \( \phi(\mathcal{D}) \) (see Amballa et al. [2021] for a definition). Theorem 1 of Amballa et al. [2021] now implies that 2) holds.

To prove that 2) implies 1), suppose \( p_{\min} = p_1 \leq p_2 \leq \cdots \leq p_f = p_{\max} \) satisfy (11). Define \( p \) as in Lemma D.1. The Lagrange polynomials defined in Lemma D.1 satisfy

\[
\begin{align*}
  l_i(p_i, p) &= 1, \quad i = 1, \ldots, f, \quad (31) \\
  l_i(p_j, p) &= 0, \quad i, j = 1, \ldots, f, \quad i \neq j, \quad (32) \\
  \frac{dl_i}{ds}(p_i, p) &= 0, \quad i = 2, \ldots, f - 1, \quad (33)
\end{align*}
\]

Equations (31), (32) and the inequalities in (34) follow by substituting appropriate values for \( s \) in (27), while (33) follows by differentiating (27) with respect to \( s \), substituting appropriately for \( s \), and then using (11).

Next, define the function \( G : \mathcal{D} \to \mathbb{R} \) by \( G(s) \stackrel{\text{def}}{=} l_1^2(s, p) + \cdots + l_f^2(s, p) - 1 \). We claim that \( G(s) \leq 0 \) for all \( s \in \mathcal{D} \). In light of Lemma D.1 and the definition of \((1,f)\)-volumetric points, our claim implies that 1) holds. Hence, to complete the proof, it is sufficient to prove our claim.

To prove our claim, note that \( G \) is a polynomial of degree \( 2(f - 1) \). Also, we observe from (31), (33) and (34) that \( p_1 \) and \( p_f \) are roots of \( G \) of multiplicity 1, while each \( p_i \) is a root of \( G \) of multiplicity at least 2 for \( i \neq 1, f \). Thus, the polynomial \( H(s) \stackrel{\text{def}}{=} (s - p_1)(s - p_2)^2 \cdots (s - p_{f-1})^2(s - p_f) \) divides \( G \). Since \( H \) also clearly has degree \( 2(f - 1) \), it follows that there exists \( K \in \mathbb{R} \) such that \( G(s) = KH(s) \) for all \( s \in \mathcal{D} \). The value of \( K \) may be computed as \( K = G'(p_1) / H'(p_1) \), where \('\) indicates the derivative. It is easy to use (31) and (32) to verify that \( G'(p_1) = 2 \frac{dl_1}{ds}(p_1, p) \), which is negative by (34). An easy calculation also yields \( H'(p_1) = (p_1 - p_2)^2 \cdots (p_1 - p_{f-1})^2(p_1 - p_f) \) which is negative since \( p_1 < p_f \). These arguments show that \( K > 0 \). Our claim now follows by noting that \( H \) takes only non-positive values on \( \mathcal{D} \). This completes the proof. \( \square \)
Algorithm 1 VSBAI-Poly: Best Arm Identification for Polynomial Rewards

1: Input: $\varepsilon > 0$, $\delta \in (0, 1)$, sub-Gaussianity parameter $\sigma$,
   $(1, f)$-volumetric points $p_1, \ldots, p_f$ for $(\phi, D)$
2: Set $B_{1,f} = [\phi(p_1), \ldots, \phi(p_f)]$
3: Initialize $k \leftarrow 1$, $r \leftarrow 0$
4: Set STOP = False
5: while STOP == False do
6:   $\bar{y}^k = []$
7:   for $t = 1, \ldots, f$ do
8:      $y_{(k-1)m+t} = g_{\mu}(p_t) + \eta_t$
9:      $\bar{y}^k \leftarrow [(\bar{y}^k)^T; y_{(k-1)f+t}]^T$
10: end for
11: $r = r + \bar{y}^k$
12: if $\beta(k, 4\varepsilon^4) < \delta$ then
13:   STOP = True
14: else
15:   $k = k + 1$
16: end if
17: end while
18: $r^* = kd$
19: $\hat{\mu}_{r^*} = \frac{1}{k} B_{1,f}^{-T} r$
20: $\hat{s} = \text{global\_optimizer}(\hat{\mu}_{r^*}, \hat{s}_{min}, \hat{s}_{max})$
21: $D_{r^*} = \text{get\_dtau}(\hat{\mu}_{r^*}, \hat{s}, \hat{s}_{min}, \hat{s}_{max}, \varepsilon)$
22: Output: $D_{r^*}$

23: Function global\_optimizer($\hat{\mu}_{r^*}, \hat{s}_{min}, \hat{s}_{max}$)
24: $\hat{\mu}'_{r^*} = \text{differentiate}(\hat{\mu}_{r^*})$
25: roots = find\_roots($\hat{\mu}'_{r^*}$)
26: roots.add($\hat{s}_{min}, \hat{s}_{max}$)
27: values = $g_{\hat{\mu}_{r^*}}(\text{roots})$
28: opt\_value = argmax(values)
29: return opt\_value

30: Function get\_dtau($\hat{\mu}_{r^*}, \hat{s}, \hat{s}_{min}, \hat{s}_{max}, \varepsilon$)
31: $d\_tau = []$
32: find\_roots($g_{\hat{\mu}_{r^*}}(\hat{s}) - g_{\hat{\mu}_{r^*}}(\hat{s}) + \varepsilon/2$)
33: roots.add($\hat{s}_{min}, \hat{s}_{max}$)
34: roots = sort(roots)
35: root\_left = get\_closest\_left\_root\_to\_\hat{s}(roots, $\hat{s}$)
36: root\_right = get\_closest\_right\_root\_to\_\hat{s}(roots, $\hat{s}$)
37: $d\_tau.add(\text{root\_left, root\_right})$
38: $d\_tau.add(\text{every\_pair\_to\_the\_left\_of\_root\_left})$
39: $d\_tau.add(\text{every\_pair\_to\_the\_right\_of\_root\_right})$
40: return $d\_tau$
Figure 3: 10 arm setting when the angles $\phi$ of the arms (3 to 10) are sampled from $\mathcal{N}(0, .09)$. $(\epsilon, \delta) = (0.1, 0.05)$ for the VSBAI algorithm.

Figure 4: 10 arm setting when the angles $\phi$ of the arms (3 to 10) are sampled uniformly from $[0, 0.1]$. $(\epsilon, \delta) = (0.1, 0.05)$ for the VSBAI algorithm.
We consider the setting outlined in subsection 6.1 and present results for a different configuration of the problem instances. We first note that the implementation of the baseline algorithms presented in Fiez et al. [2019], Jedra and Proutiere [2020a], and Soare et al. [2014] for the setting in subsection 6.1 is true when the angles $\phi_i$ for $i = 3, \ldots, n$ are sampled from a uniform distribution $[0, 0.1]$ rather than a Gaussian distribution as in subsection 6.1. We therefore present experimental results for this uniform setting and provide a comparison of sample complexity and run time as in subsection 6.1. Note that we are able reproduce the results reported in Jedra and Proutiere [2020b] (see Table 2 and Table 3 of Jedra and Proutiere [2020b]).

We observe from tables 3 that the sample complexity of VSBAI is greater than the other baselines. However, we argue that the instances generated in this setting are simple and in situations where it is difficult to separate out the best-arm from the next best (like when the angles $\phi_i$ of the arms are sampled from Gaussian Gaussian setting in 6.1), all these baselines suffer from huge sample complexities and run-times. In other words VSBAI is independent of the way the instances are generated but on the other hand all the other baselines are not robust, hence can potentially perform badly in adversarial environments. Table 4 gives a comparison of the run-times for this setting. The results shown are obtained after averaging over 20 seeds.
### Table 3: Expected sample complexity for the setting described in Appendix G

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>LazyTS Mean</th>
<th>Std</th>
<th>Rage Mean</th>
<th>Std</th>
<th>Oracle Mean</th>
<th>Std</th>
<th>VSBAI Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>335.1</td>
<td>22.71</td>
<td>524.1</td>
<td>33.84</td>
<td>347.1</td>
<td>32.22</td>
<td>47693.4</td>
<td>105.32</td>
</tr>
<tr>
<td>20</td>
<td>423.05</td>
<td>28.92</td>
<td>683.05</td>
<td>92.07</td>
<td>356.15</td>
<td>32.31</td>
<td>47424.1</td>
<td>41.32</td>
</tr>
<tr>
<td>100</td>
<td>421.75</td>
<td>32.11</td>
<td>1038.75</td>
<td>148.01</td>
<td>426.4</td>
<td>39.76</td>
<td>47271.4</td>
<td>13.49</td>
</tr>
<tr>
<td>1000</td>
<td>419.65</td>
<td>28.43</td>
<td>1152.7</td>
<td>50.23</td>
<td>476.45</td>
<td>40.74</td>
<td>47222.7</td>
<td>1.27</td>
</tr>
<tr>
<td>2500</td>
<td>446.15</td>
<td>29.06</td>
<td>1447.3</td>
<td>150.46</td>
<td>481.8</td>
<td>41.33</td>
<td>47219.8</td>
<td>0.37</td>
</tr>
<tr>
<td>5000</td>
<td>431.65</td>
<td>32.23</td>
<td>1546.9</td>
<td>160.17</td>
<td>510.05</td>
<td>48.87</td>
<td>47219.9</td>
<td>0.41</td>
</tr>
</tbody>
</table>

### Table 4: Run-time in seconds for the setting described in Appendix G

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>LazyTS Mean</th>
<th>Std</th>
<th>Rage Mean</th>
<th>Std</th>
<th>Oracle Mean</th>
<th>Std</th>
<th>VSBAI Mean</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.27</td>
<td>0.01</td>
<td>0.05</td>
<td>0.001</td>
<td>0.001</td>
<td>0.0</td>
<td>1.30</td>
<td>0.02</td>
</tr>
<tr>
<td>20</td>
<td>0.33</td>
<td>0.02</td>
<td>0.06</td>
<td>0.00</td>
<td>0.001</td>
<td>0.0</td>
<td>1.39</td>
<td>0.04</td>
</tr>
<tr>
<td>100</td>
<td>0.39</td>
<td>0.02</td>
<td>1.09</td>
<td>0.07</td>
<td>0.02</td>
<td>0.0</td>
<td>1.41</td>
<td>0.04</td>
</tr>
<tr>
<td>1000</td>
<td>34.78</td>
<td>1.06</td>
<td>27.61</td>
<td>0.33</td>
<td>0.69</td>
<td>0.02</td>
<td>1.44</td>
<td>0.04</td>
</tr>
<tr>
<td>2500</td>
<td>211.24</td>
<td>8.22</td>
<td>335.82</td>
<td>2.46</td>
<td>0.65</td>
<td>0.02</td>
<td>1.56</td>
<td>0.04</td>
</tr>
<tr>
<td>5000</td>
<td>422.35</td>
<td>22.45</td>
<td>884.32</td>
<td>3.85</td>
<td>0.89</td>
<td>0.03</td>
<td>2.17</td>
<td>0.03</td>
</tr>
</tbody>
</table>
References


