Truncated Variance Reduction: A Unified Approach to Bayesian Optimization and Level-Set Estimation

Ilija Bogunovic[1], Jonathan Scarlett[1], Andreas Krause[2], Volkan Cevher[1]
[1] Laboratory for Information and Inference Systems (LIONS), EPFL
[2] Learning and Adaptive Systems Group, ETH Zürich

{ilija.bogunovic,jonathan.scarlett,volkan.cevher}@epfl.ch, krausea@ethz.ch

Abstract

We present a new algorithm, truncated variance reduction (TRUVAR), that treats Bayesian optimization (BO) and level-set estimation (LSE) with Gaussian processes in a unified fashion. The algorithm greedily shrinks a sum of truncated variances within a set of potential maximizers (BO) or unclassified points (LSE), which is updated based on confidence bounds. TRUVAR is effective in several important settings that are typically non-trivial to incorporate into myopic algorithms, including pointwise costs and heteroscedastic noise. We provide a general theoretical guarantee for TRUVAR covering these aspects, and use it to recover and strengthen existing results on BO and LSE. Moreover, we provide a new result for a setting where one can select from a number of noise levels having associated costs. We demonstrate the effectiveness of the algorithm on both synthetic and real-world data sets.

1   Introduction

Bayesian optimization (BO) [1] provides a powerful framework for automating design problems, and finds applications in robotics, environmental monitoring, and automated machine learning, just to name a few. One seeks to find the maximum of an unknown reward function that is expensive to evaluate, based on a sequence of suitably-chosen points and noisy observations. Numerous BO algorithms have been presented previously; see Section 1.1 for an overview.

Level-set estimation (LSE) [2] is closely related to BO, with the added twist that instead of seeking a maximizer, one seeks to classify the domain into points that lie above or below a certain threshold. This is of considerable interest in applications such as environmental monitoring and sensor networks, allowing one to find all "sufficiently good" points rather than the best point alone.

While BO and LSE are closely related, they are typically studied in isolation. In this paper, we provide a unified treatment of the two via a new algorithm, *Truncated Variance Reduction* (TRUVAR), which enjoys theoretical guarantees, good computational complexity, and the versatility to handle important settings such as pointwise costs, non-constant noise, and multi-task scenarios. The main result of this paper applies to the former two settings, and even the fixed-noise and unit-cost case, we refine existing bounds via a significantly improved dependence on the noise level.

1.1   Previous Work

Three popular myopic techniques for Bayesian optimization are expected improvement (EI), probability of improvement (PI), and Gaussian process upper confidence bound (GP-UCB) [1, 3], each of which chooses the point maximizing an acquisition function depending directly on the current posterior mean and variance. In [4], the GP-UCB-PE algorithm was presented for BO, choosing the highest-variance point within a set of potential maximizers that is updated based on confidence bounds. Another relevant BO algorithm is BaMSOO [5], which also keeps track of potential maximizers, but instead chooses points based on a global optimization technique called simultaneous

online optimization (SOO). An algorithm for level-set estimation with GPs is given in [2], which keeps track of a set of unclassified points. These algorithms are computationally efficient and have various theoretical guarantees, but it is unclear how best to incorporate aspects such as pointwise costs and heteroscedastic noise [6]. The same is true for the Straddle heuristic for LSE [7].

Entropy search (ES) [8] and its predictive version [9] choose points to reduce the uncertainty of the location of the maximum, doing so via a *one-step lookahead* of the posterior rather than only the current posterior. While this is more computationally expensive, it also permits versatility with respect to costs [6], heteroscedastic noise [10], and multi-task scenarios [6]. A recent approach called minimum regret search (MRS) [11] also performs a look-ahead, but instead chooses points to minimize the regret. To our knowledge, no theoretical guarantees have been provided for these.

The multi-armed bandit (MAB) [12] literature has developed alongside the BO literature, with the two often bearing similar concepts. The MAB literature is far too extensive to cover here, but we briefly mention some variants relevant to this paper. Extensive attention has been paid to the *best-arm identification* problem [13], and cost constraints have been incorporated in a variety of forms [14]. Moreover, the concept of "zooming in" to the optimal point has been explored [15]. In general, the assumptions and analysis techniques in the MAB and BO literature are quite different.

1.2 Contributions

We present a unified analysis of Bayesian optimization and level-set estimation via a new algorithm Truncated Variance Reduction (TRUVAR). The algorithm works by keeping track of a set of potential maximizers (BO) or unclassified points (LSE), selecting points that shrink the uncertainty within that set up to a truncation threshold, and updating the set using confidence bounds. Similarly to ES and MRS, the algorithm performs a one-step lookahead that is highly beneficial in terms of versatility. However, unlike these previous works, our lookahead avoids the computationally expensive task of averaging over the posterior distribution and the observations.

Also in contrast with ES and MRS, we provide theoretical bounds for TRUVAR characterizing the cost required to achieve a certain accuracy in finding a near-optimal point (BO) or in classifying each point in the domain (LSE). By applying this to the standard BO setting, we not only recover existing results [2, 4], but we also strengthen them via a significantly improved dependence on the noise level, with better asymptotics in the small noise limit. Moreover, we provide a novel result for a setting in which the algorithm can choose the noise level, each coming with an associated cost.

Finally, we compare our algorithm to previous works on several synthetic and real-world data sets, observing it to perform favorably in a variety of settings.

2 Problem Setup and Proposed Algorithm

Setup: We seek to sequentially optimize an unknown reward function $f(x)$ over a finite domain D.[1] At time t, we query a single point $x_t \in D$ and observe a noisy sample $y_t = f(x_t) + z_t$, where $z_t \sim N(0, \sigma^2(x_t))$ for some known noise function $\sigma^2(\cdot) : D \to \mathbb{R}_+$. Thus, in general, some points may be noisier than others, in which case we have *heteroscedastic noise* [10]. We associate with each point a *cost* according to some known cost function $c : D \to \mathbb{R}_+$. If both $\sigma^2(\cdot)$ and $c(\cdot)$ are set to be constant, then we recover the standard homoscedastic and unit-cost setting.

We model $f(x)$ as a Gaussian process (GP) [16] having mean zero and kernel function $k(x, x')$, normalized so that $k(x, x) = 1$ for all $x \in D$. The posterior distribution of f given the points and observations up to time t is again a GP, with the posterior mean and variance given by [10]

$$\mu_t(x) = \mathbf{k}_t(x)^T (\mathbf{K}_t + \mathbf{\Sigma}_t)^{-1} \mathbf{y}_t \tag{1}$$

$$\sigma_t(x)^2 = k(x, x) - \mathbf{k}_t(x)^T (\mathbf{K}_t + \mathbf{\Sigma}_t)^{-1} \mathbf{k}_t(x), \tag{2}$$

where $\mathbf{k}_t(x) = [k(x_i, x)]_{i=1}^t$, $\mathbf{K}_t = [k(x_t, x_{t'})]_{t,t'}$, and $\mathbf{\Sigma}_t = \mathrm{diag}(\sigma^2(x_1), \ldots, \sigma^2(x_t))$. We also let $\sigma_{t-1|x}^2(\overline{x})$ denote the posterior variance of $\overline{x}$ upon observing x along with $x_1, \cdots, x_{t-1}$.

---

[1]Extensions to continuous domains are discussed in the supplementary material.

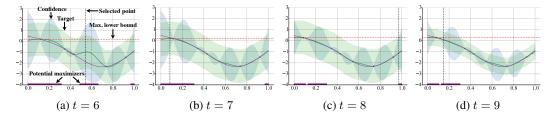| (a) $t = 6$ | (b) $t = 7$ | (c) $t = 8$ | (d) $t = 9$ |

Figure 1: An illustration of the TRUVAR algorithm. In (a), (b), and (c), three points within the set of potential maximizers M_t are selected in order to bring the confidence bounds to within the target range, and M_t shrinks during this process. In (d), the target confidence width shrinks as a result of the last selected point bringing the confidence within M_t to within the previous target.

We consider both *Bayesian optimization*, which consists of finding a point whose function value is as high as possible, and *level-set estimation*, which consists of classifying the domain according into points that lie above or below a given threshold h. The precise performance criteria for these settings are given in Definition 3.1 below. Essentially, after spending a certain cost we report a point (BO) or a classification (LSE), but there is no preference on the values of $f(x_t)$ for the points x_t chosen before coming to such a decision (in contrast with other notions such as cumulative regret).

TRUVAR algorithm: Our algorithm is described in Algorithm 1, making use of the updates described in Algorithm 2. The algorithm keeps track of a sequence of unclassified points M_t, representing potential maximizers for BO or points close to h for LSE. This set is updated based on the confidence bounds depending on constants $\beta_{(i)}$. The algorithm proceeds in epochs, where in the i-th epoch it seeks to bring the confidence $\beta_{(i)}^{1/2}\sigma_t(x)$ of points within M_t below a target value $\eta_{(i)}$. It does this by greedily minimizing the sum of truncated variances $\sum_{\overline{x} \in M_{t-1}} \max\{\beta_{(i)}\sigma_{t-1|x}^2(\overline{x}), \eta_{(i)}\}$ arising from choosing the point x, along with a normalization and division by $c(x)$ to favor low-cost points. The truncation by $\eta_{(i)}$ in this decision rule means that once the confidence of a point is below the current target value, there is no preference in making it any lower (until the target is decreased). Once the confidence of every point in M_t is less than a factor $1 + \overline{\delta}$ above the target value, the target confidence is reduced according to a multiplication by $r \in (0, 1)$. An illustration of the process is given in Figure 1, with details in the caption.

For level-set estimation, we also keep track of the sets H_t and L_t, containing points believed to have function values above and below h, respectively. The constraint $x \in M_{t-1}$ in (5)–(7) ensures that $\{M_t\}$ is non-increasing with respect to inclusion, and H_t and L_t are non-decreasing.

Algorithm 1 Truncated Variance Reduction (TRUVAR)

Input: Domain D, GP prior (μ_0, σ_0, k), confidence bound parameters $\overline{\delta} > 0$, $r \in (0, 1)$, $\{\beta_{(i)}\}_{i \geq 1}$, $\eta_{(1)} > 0$, and for LSE, level-set threshold h
1: Initialize the epoch number $i = 1$ and potential maximizers $M_{(0)} = D$.
2: **for** $t = 1, 2, \ldots$ **do**
3:     Choose

$$x_t = \arg\max_{x \in D} \frac{\sum_{\overline{x} \in M_{t-1}} \max\{\beta_{(i)}\sigma_{t-1}^2(\overline{x}), \eta_{(i)}^2\} - \sum_{\overline{x} \in M_{t-1}} \max\{\beta_{(i)}\sigma_{t-1|x}^2(\overline{x}), \eta_{(i)}^2\}}{c(x)}. \quad (3)$$

4:     Observe the noisy function sample y_t, and update according to Algorithm 2 to obtain M_t, μ_t, σ_t, l_t and u_t, as well as H_t and L_t in the case of LSE
5:     **while** $\max_{x \in M_t} \beta_{(i)}^{1/2}\sigma_t(x) \leq (1 + \overline{\delta})\eta_{(i)}$ **do**
6:         Increment i, set $\eta_{(i)} = r \times \eta_{(i-1)}$.

The choices of $\beta_{(i)}$, $\overline{\delta}$, and r are discussed in Section 4. As with previous works, the kernel is assumed known in our theoretical results, whereas in practice it is typically learned from training data [3]. Characterizing the effect of model mismatch or online hyperparameter updates is beyond the scope of this paper, but is an interesting direction for future work.

Algorithm 2 Parameter Updates for TRUVAR

Input: Selected points and observations $\{x_{t'}\}_{t'=1}^{t}$; $\{y_{t'}\}_{t'=1}^{t}$, previous sets M_{t-1}, H_{t-1}, L_{t-1}, parameter $\beta_{(i)}^{1/2}$, and for LSE, level-set threshold h.

1: Update μ_t and σ_t according to (1)–(2), and form the upper and lower confidence bounds

$$u_t(x) = \mu_t(x) + \beta_{(i)}^{1/2}\sigma_t(x), \quad \ell_t(x) = \mu_t(x) - \beta_{(i)}^{1/2}\sigma_t(x). \tag{4}$$

2: For BO, set

$$M_t = \left\{ x \in M_{t-1} : u_t(x) \geq \max_{\overline{x} \in M_{t-1}} \ell_t(\overline{x}) \right\}, \tag{5}$$

or for LSE, set

$$M_t = \left\{ x \in M_{t-1} : u_t(x) \geq h \text{ and } \ell_t(x) \leq h \right\} \tag{6}$$

$$H_t = H_{t-1} \cup \left\{ x \in M_{t-1} : \ell_t(x) > h \right\}, \quad L_t = L_{t-1} \cup \left\{ x \in M_{t-1} : u_t(x) < h \right\}. \tag{7}$$

Some variants of our algorithm and theory are discussed in the supplementary material due to lack of space, including *pure* variance reduction, non-Bayesian settings [3], continuous domains [3], the batch setting [4], and implicit thresholds for level-set estimation [2].

3 Theoretical Bounds

In order to state our results for BO and LSE in a unified fashion, we define a notion of ϵ-*accuracy* for the two settings. That is, we define this term differently in the two scenarios, but then we provide theorems that simultaneously apply to both. All proofs are given in the supplementary material.

Definition 3.1. *After time step t of* TRUVAR*, we use the following terminology:*

- *For BO, the set M_t is ϵ-accurate if it contains all true maxima $x^* \in \arg\max_x f(x)$, and all of its points satisfy $f(x^*) - f(x) \leq \epsilon$.*

- *For LSE, the triplet (M_t, H_t, L_t) is ϵ-accurate if all points in H_t satisfy $f(x) > h$, all points in L_t satisfy $f(x) < h$, and all points in M_t satisfy $|f(x) - h| \leq \frac{\epsilon}{2}$.*

In both cases, the cumulative cost *after time t is defined as $C_t = \sum_{t'=1}^{t} c(x_{t'})$.*

We use $\frac{\epsilon}{2}$ in the LSE setting instead of ϵ since this creates a region of size ϵ where the function value lies, which is consistent with the BO setting. Our performance criterion for level-set estimation is slightly different from that of [2], but the two are closely related.

3.1 General Result

Preliminary definitions: Suppose that the $\{\beta_{(i)}\}$ are chosen to ensure valid confidence bounds, i.e., $\ell_t(x) \leq f(x) \leq u_t(x)$ with high probability; see Theorem 3.1 and its proof below for such choices. In this case, we have after the i-th epoch that all points are either already discarded (BO) or classified (LSE), or are known up to the confidence level $(1 + \overline{\delta})\eta_{(i)}$. For the points with such confidence, we have $u_t(x) - \ell_t(x) \leq 2(1 + \overline{\delta})\eta_{(i)}$, and hence

$$u_t(x) \leq \ell_t(x) + 2(1 + \overline{\delta})\eta_{(i)} \leq f(x) + 2(1 + \overline{\delta})\eta_{(i)}, \tag{8}$$

and similarly $\ell_t(x) \geq f(x) - 2(1 + \overline{\delta})\eta_{(i)}$. This means that all points other than those within a gap of width $4(1 + \overline{\delta})\eta_{(i)}$ must have been discarded or classified:

$$M_t \subseteq \left\{ x : f(x) \geq f(x^*) - 4(1 + \overline{\delta})\eta_{(i)} \right\} =: \overline{M}_{(i)} \quad \text{(BO)} \tag{9}$$

$$M_t \subseteq \left\{ x : |f(x) - h| \leq 2(1 + \overline{\delta})\eta_{(i)} \right\} =: \overline{M}_{(i)} \quad \text{(LSE)} \tag{10}$$

Since no points are discarded or classified initially, we define $\overline{M}_{(0)} = D$.

4

For a collection of points $S = (x'_1, \ldots, x'_{|S|})$, possibly containing duplicates, we write the total cost as $c(S) = \sum_{i=1}^{|S|} c(x'_i)$. Moreover, we denote the posterior variance upon observing the points up to time $t-1$ *and* the additional points in S by $\sigma_{t-1|S}(\overline{x})$. Therefore, $c(x) = c(\{x\})$ and $\sigma_{t-1|x}(\overline{x}) = \sigma_{t-1|\{x\}}(\overline{x})$. The minimum cost (respectively, maximum cost) is denoted by $c_{\min} = \min_{x \in D} c(x)$ (respectively, $c_{\max} = \max_{x \in D} c(x)$).

Finally, we introduce the quantity

$$C^*(\xi, M) = \min_S \left\{ c(S) \; : \; \max_{\overline{x} \in M} \sigma_{0|S}(\overline{x}) \le \xi \right\}, \tag{11}$$

representing the minimum cost to achieve a posterior standard deviation of at most ξ within M.

Main result: In all of our results, we make the following assumption.

Assumption 3.1. *The kernel $k(x, x')$ is such that the variance reduction function*

$$\psi_{t,x}(S) = \sigma_t^2(x) - \sigma_{t|S}^2(x) \tag{12}$$

is submodular [17] for any time t, and any selected points $(x_1, \ldots, x_t)$ and query point x.

This assumption has been used in several previous works based on Gaussian processes, and sufficient conditions for its validity can be found in [18, Sec. 8]. We now state the following general guarantee.

Theorem 3.1. *Fix $\epsilon > 0$ and $\delta \in (0, 1)$, and suppose there exist values $\{C_{(i)}\}$ and $\{\beta_{(i)}\}$ such that*

$$C_{(i)} \ge C^* \left(\frac{\eta_{(i)}}{\beta_{(i)}^{1/2}}, \overline{M}_{(i-1)} \right) \log \frac{|\overline{M}_{(i-1)}|\beta_{(i)}}{\overline{\delta}^2 \eta_{(i)}^2} + c_{\max}, \tag{13}$$

and

$$\beta_{(i)} \ge 2 \log \frac{|D| \left(\sum_{i' \le i} C_{(i')} \right)^2 \pi^2}{6\delta c_{\min}^2}. \tag{14}$$

Then if TRUVAR is run with these choices of $\beta_{(i)}$ until the cumulative cost reaches

$$C_\epsilon = \sum_{i \,:\, 4(1+\overline{\delta})\eta_{(i-1)} > \epsilon} C_{(i)}, \tag{15}$$

then with probability at least $1 - \delta$, we have ϵ-accuracy.

While this theorem is somewhat abstract, it captures the fact that the algorithm improves when points having a lower cost and/or lower noise are available, since both of these lead to a smaller value of $C^*(\xi, M)$; the former by directly incurring a smaller cost, and the latter by shrinking the variance more rapidly. Below, we apply this result to some important cases.

3.2 Results for Specific Settings

Homoscedastic and unit-cost setting: Define the maximum mutual information [3]

$$\gamma_T = \max_{x_1, \ldots, x_T} \frac{1}{2} \log \det \left(\mathbf{I}_T + \sigma^{-2} \mathbf{K}_T \right), \tag{16}$$

and consider the case that $\sigma^2(x) = \sigma^2$ and $c(x) = 1$. In the supplementary material, we provide a theorem with a condition for ϵ-accuracy of the form $T \ge \Omega^* \left(\frac{C_1 \gamma_T \beta_T}{\epsilon^2} + 1 \right)$ with $C_1 = \frac{1}{\log(1+\sigma^{-2})}$, thus matching [2, 4] up to logarithmic factors. In the following, we present a refined version that has a significantly better dependence on the noise level, thus exemplifying that a more careful analysis of (13) can provide improvements over the standard bounding techniques.

Corollary 3.1. *Fix $\epsilon > 0$ and $\delta \in (0, 1)$, define $\beta_T = 2 \log \frac{|D|T^2 \pi^2}{6\delta}$, and set $\eta_{(1)} = 1$ and $r = \frac{1}{2}$. There exist choices of $\beta_{(i)}$ (not depending on the time horizon T) such that we have ϵ-accuracy with probability at least $1 - \delta$ once the following condition holds:*

$$T \ge \left(2\sigma^2 \gamma_T \beta_T \frac{96(1+\overline{\delta})^2}{\epsilon^2} + C_1 \gamma_T \beta_T \frac{6(1+\overline{\delta})^2}{\sigma^2} + 2 \left\lceil \log_2 \frac{32(1+\overline{\delta})^2}{\epsilon \sigma} \right\rceil \right) \log \frac{16(1+\overline{\delta})^2 |D| \beta_T}{\overline{\delta}^2 \epsilon^2}, \tag{17}$$

where $C_1 = \frac{1}{\log(1+\sigma^{-2})}$. This condition is of the form $T \ge \Omega^* \left(\frac{\sigma^2 \gamma_T \beta_T}{\epsilon^2} + \frac{C_1 \gamma_T \beta_T}{\sigma^2} + 1 \right)$.

The choices $\eta_{(1)} = 1$ and $r = \frac{1}{2}$ are made for mathematical convenience, and a similar result follows for any other choices $\eta_{(1)} > 0$ and $r \in (0, 1)$, possibly with different constant factors.

As $\sigma^2 \to \infty$ (i.e., high noise), both of the above-mentioned bounds have noise dependence $O^*(\sigma^2)$, since $\log(1 + \alpha^{-1}) = O(\alpha^{-1})$ as $\alpha \to \infty$. On the other hand, as $\sigma^2 \to 0$ (i.e., low noise), C_1 is logarithmic, and Corollary 3.1 is significantly better provided that $\epsilon \ll \sigma$.

Choosing the noise and cost: Here we consider the setting that there is a domain of points D_0 that the reward function depends on, and alongside each point we can *choose* a noise variance $\sigma^2(k)$ ($k = 1, \ldots, K$). Hence, $D = D_0 \times \{1, \cdots, K\}$. Lower noise variances incur a higher cost according to a cost function $c(k)$.

Corollary 3.2. *For each $k = 1, \cdots, K$, let $T^*(k)$ denote the smallest value of T such that* (17) *holds with $\sigma^2(k)$ in place of σ^2, and with $\beta_T = 2 \log \frac{|D|T^2 c_{\max}^2 \pi^2}{6\delta c_{\min}^2}$. Then, under the preceding setting, there exist choices of $\beta_{(i)}$ (not depending on T) such that we have ϵ-accuracy with probability at least $1-\delta$ once the cumulative cost reaches $\min_k c(k)T^*(k)$.*

This result roughly states that we obtain a bound as good as that obtained by sticking to *any fixed choice of noise level*. In other words, every choice of noise (and corresponding cost) corresponds to a different version of a BO or LSE algorithm (e.g., [2, 4]), and our algorithm has a similar performance guarantee to the best among all of those. This is potentially useful in avoiding the need for running an algorithm once per noise level and then choosing the best-performing one. Moreover, we found numerically that beyond matching the best fixed noise strategy, we can strictly improve over it by mixing the noise levels; see Section 4.

4 Experimental Results

We evaluate our algorithm in both the level-set estimation and Bayesian optimization settings.

Parameter choices: As with previous GP-based algorithms that use confidence bounds, our theoretical choice of $\beta_{(i)}$ in TRUVAR is typically overly conservative. Therefore, instead of using (14) directly, we use a more aggressive variant with similar dependence on the domain size and time: $\beta_{(i)} = a \log(|D|t_{(i)}^2)$, where $t_{(i)}$ is the time at which the epoch starts, and a is a constant. Instead of the choice $a = 2$ dictated by (14), we set $a = 0.5$ for BO to avoid over-exploration. We found exploration to be slightly more beneficial for LSE, and hence set $a = 1$ for this setting. We found TRUVAR to be quite robust with respect to the choices of the remaining parameters, and simply set $\eta_{(1)} = 1$, $r = 0.1$, and $\overline{\delta} = 0$ in all experiments; while our theory assumes $\overline{\delta} > 0$, in practice there is negligible difference between choosing zero and a small positive value.

Level-set estimation: For the LSE experiments, we use a common classification rule in all algorithms, classifying the points according to the posterior mean as $\hat{H}_t = \{x : \mu_t(x) \geq h\}$ and $\hat{L}_t = \{x : \mu_t(x) < h\}$. The classification accuracy is measured by the F_1-score (i.e., the harmonic mean of precision and recall) with respect to the true super- and sub-level sets.

We compare TRUVAR against the GP-based LSE algorithm [2], which we name via the authors' surnames as GCHK, as well as the state-of-the-art straddle (STR) heuristic [7] and the maximum variance rule (VAR) [2]. Descriptions can be found in the supplementary material. GCHK includes an exploration constant β_t, and we follow the recommendation in [2] of setting $\beta_t^{1/2} = 3$.

Lake data (unit cost): We begin with a data set from the domain of environmental monitoring of inland waters, consisting of 2024 in situ measurements of chlorophyll concentration within a vertical transect plane, collected by an autonomous surface vessel in Lake Zürich [19]. As in [2], our goal is to detect regions of high concentration. We evaluate each algorithm on a 50×50 grid of points, with the corresponding values coming from the GP posterior that was derived using the original data (see Figure 2d). We use the Matérn-5/2 ARD kernel, setting its hyperparameters by maximizing the likelihood on the second (smaller) available dataset. The level-set threshold h is set to 1.5.

In Figure 2a, we show the performance of the algorithms averaged over 100 different runs; here the randomness is only with respect to the starting point, as we are in the noiseless setting. We observe that in this unit-cost case, TRUVAR performs similarly to GCHK and STR. All three methods outperform VAR, which is good for global exploration but less suited to level-set estimation.
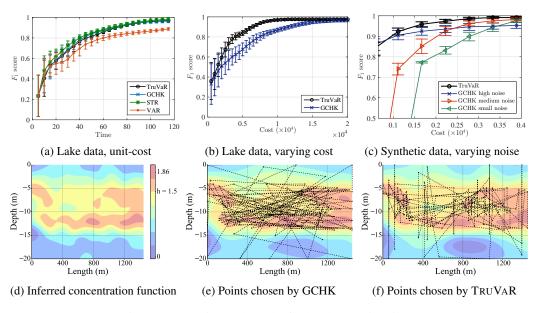
(a) Lake data, unit-cost (b) Lake data, varying cost (c) Synthetic data, varying noise

(d) Inferred concentration function (e) Points chosen by GCHK (f) Points chosen by TRUVAR

Figure 2: Experimental results for level-set estimation.

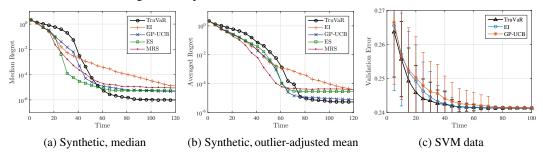(a) Synthetic, median (b) Synthetic, outlier-adjusted mean (c) SVM data

Figure 3: Experimental results for Bayesian optimization.

Lake data (varying cost): Next, we modify the above setting by introducing pointwise costs that are a function of the previous sampled point x', namely, $c_{x'}(x) = 0.25|x_1 - x'_1| + 4(|x_2| + 1)$, where x_1 is the vessel position and x_2 is the depth. Although we did not permit such a dependence on x' in our original setup, the algorithm itself remains unchanged. Our choice of cost penalizes the distance traveled $|x_1 - x'_1|$, as well as the depth of the measurement $|x_2|$. Since incorporating costs into existing algorithms is non-trivial, we only compare against the original version of GCHK that ignores costs.

In Figure 2b, we see that TruVaR significantly outperforms GCHK, achieving a higher F_1 score for a significantly smaller cost. The intuition behind this can be seen in Figures 2e and 2f, where we show the points sampled by TruVaR and GCHK in one experiment run, connecting all pairs of consecutive points. GCHK is designed to pick few points, but since it ignores costs, the distance traveled is large. In contrast, by incorporating costs, TRUVAR tends to travel small distances, often even staying in the same x_1 location to take measurements at multiple depths x_2.

Synthetic data with multiple noise levels: In this experiment, we demonstrate Corollary 3.2 by considering the setting in which the algorithm can choose the sampling noise variance and incur the associated cost. We use a synthetic function sampled from a GP on a 50×50 grid with an isotropic squared exponential kernel having length scale $l = 0.1$ and unit variance, and set $h = 2.25$. We use three different noise levels, $\sigma^2 \in \{10^{-6}, 10^{-3}, 0.05\}$, with corresponding costs $\{15, 10, 2\}$.

We run GCHK separately for each of the three noise levels, while running TRUVAR as normal and allowing it to mix between the noise levels. The resulting F_1-scores are shown in Figure 2c. The best-performing version of GCHK changes throughout the time horizon, while TRUVAR is consistently better than all three. A discussion on how TRUVAR mixes between the noise levels can be found in the supplementary material.

Bayesian optimization. We now provide the results of two experiments for the BO setting.

Synthetic data: We first conduct a similar experiment as that in [8, 11], generating 200 different test functions defined on $[0, 1]^2$. To generate a single test function, 200 points are chosen uniformly at random from $[0, 1]^2$, their function values are generated from a GP using an isotropic squared exponential kernel with length scale $l = 0.1$ and unit variance, and the resulting posterior mean forms the function on the whole domain $[0, 1]^2$. We subsequently assume that samples of this function are corrupted by Gaussian noise with $\sigma^2 = 10^{-6}$. The extension of TRUVAR to continuous domains is straightforward, and is explained in the supplementary material. For all algorithms considered, we evaluate the performance according to the regret of a single reported point, namely, the one having the highest posterior mean.

We compare the performance of TRUVAR against expected improvement (EI), GP-upper confidence bound (GP-UCB), entropy search (ES) and minimum regret search (MRS), whose acquisition functions are outlined in the supplementary material. We use publicly available code for ES and MRS [20]. The exploration parameter β_t in GP-UCB is set according to the recommendation in [3] of dividing the theoretical value by five, and the parameters for ES and MRS are set according to the recommendations given in [11, Section 5.1].

Figure 3a plots the median of the regret, and Figure 3b plots the mean after removing outliers (i.e., the best and worst 5% of the runs). In the earlier rounds, ES and MRS provide the best performance, while TRUVAR improves slowly due to exploration. However, the regret of TRUVAR subsequently drops rapidly, giving the best performance in the later rounds after "zooming in" towards the maximum. GP-UCB generally performs well with the aggressive choice of β_t, despite previous works' experiments revealing it to perform poorly with the theoretical value.

Hyperparameter tuning data: In this experiment, we use the *SVM on grid* dataset, previously used in [21]. A $25 \times 14 \times 4$ grid of hyperparameter configurations resulting in 1400 data points was pre-evaluated, forming the search space. The goal is to find a configuration with small validation error. We use a Matérn-5/2 ARD kernel, and re-learn its hyperparameters by maximizing the likelihood after sampling every 3 points. Since the hyperparameters are not fixed in advance, we replace M_{t-1} by D in (5) to avoid incorrectly ruling points out early on, allowing some removed points to be added again in later steps. Once the estimated hyperparameters stop to vary significantly, the size of the set of potential maximizers decreases almost monotonically. Since we consider the noiseless setting here, we measure performance using the simple regret, i.e., the best point found so far.

We again average over 100 random starting points, and plot the resulting validation error in Figure 3c. Even in this noiseless and unit-cost setting that EI and GP-UCB are suited to, we find that TRUVAR performs slightly better, giving a better validation error with smaller error bars.

5   Conclusion

We highlight the following aspects in which TRUVAR is versatile:

- **Unified optimization and level-set estimation:** These are typically treated separately, whereas TRUVAR and its theoretical guarantees are essentially identical in both cases
- **Actions with costs:** TRUVAR naturally favors cost-effective points, as this is directly incorporated into the acquisition function.
- **Heteroscedastic noise:** TRUVAR chooses points that effectively shrink the variance of *other* points, thus directly taking advantage of situations in which some points are noisier than others.
- **Choosing the noise level:** We provided novel theoretical guarantees for the case that the algorithm can choose both a point and a noise level, *cf.*, Corollary 3.2.

Hence, TRUVAR directly handles several important aspects that are non-trivial to incorporate into myopic algorithms. Moreover, compared to other BO algorithms that perform a lookahead (e.g., ES and MRS), TRUVAR avoids the computationally expensive task of averaging over the posterior and/or measurements, and comes with rigorous theoretical guarantees.

References

[1] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, 2016.

[2] A. Gotovos, N. Casati, G. Hitz, and A. Krause, "Active learning for level set estimation," in *Int. Joint. Conf. Art. Intel.*, 2013.

[3] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Information-theoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012.

[4] E. Contal, D. Buffoni, A. Robicquet, and N. Vayatis, *Machine Learning and Knowledge Discovery in Databases*. Springer Berlin Heidelberg, 2013, ch. Parallel Gaussian Process Optimization with Upper Confidence Bound and Pure Exploration, pp. 225–240.

[5] Z. Wang, B. Shakibi, L. Jin, and N. de Freitas, "Bayesian multi-scale optimistic optimization," http://arxiv.org/abs/1402.7005.

[6] K. Swersky, J. Snoek, and R. P. Adams, "Multi-task Bayesian optimization," in *Adv. Neur. Inf. Proc. Sys. (NIPS)*, 2013, pp. 2004–2012.

[7] B. Bryan and J. G. Schneider, "Actively learning level-sets of composite functions," in *Int. Conf. Mach. Learn. (ICML)*, 2008.

[8] P. Hennig and C. J. Schuler, "Entropy search for information-efficient global optimization," *J. Mach. Learn. Research*, vol. 13, no. 1, pp. 1809–1837, 2012.

[9] J. M. Hernández-Lobato, M. W. Hoffman, and Z. Ghahramani, "Predictive entropy search for efficient global optimization of black-box functions," in *Adv. Neur. Inf. Proc. Sys. (NIPS)*, 2014, pp. 918–926.

[10] P. W. Goldberg, C. K. Williams, and C. M. Bishop, "Regression with input-dependent noise: A Gaussian process treatment," *Adv. Neur. Inf. Proc. Sys. (NIPS)*, vol. 10, pp. 493–499, 1997.

[11] J. H. Metzen, "Minimum regret search for single-and multi-task optimization," in *Int. Conf. Mach. Learn. (ICML)*, 2016.

[12] S. Bubeck and N. Cesa-Bianchi, *Regret Analysis of Stochastic and Nonstochastic Multi-Armed Bandit Problems*, ser. Found. Trend. Mach. Learn. Now Publishers, 2012.

[13] K. Jamieson and R. Nowak, "Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting," in *Ann. Conf. Inf. Sci. Sys. (CISS)*, 2014, pp. 1–6.

[14] O. Madani, D. J. Lizotte, and R. Greiner, "The budgeted multi-armed bandit problem," in *Learning Theory*. Springer, 2004, pp. 643–645.

[15] R. Kleinberg, A. Slivkins, and E. Upfal, "Multi-armed bandits in metric spaces," in *Proc. ACM Symp. Theory Comp.*, 2008.

[16] C. E. Rasmussen, "Gaussian processes for machine learning." MIT Press, 2006.

[17] A. Krause and D. Golovin, "Submodular function maximization," *Tractability: Practical Approaches to Hard Problems*, vol. 3, 2012.

[18] A. Das and D. Kempe, "Algorithms for subset selection in linear regression," in *Proc. ACM Symp. Theory Comp. (STOC)*. ACM, 2008, pp. 45–54.

[19] G. Hitz, F. Pomerleau, M.-E. Garneau, E. Pradalier, T. Posch, J. Pernthaler, and R. Y. Siegwart, "Autonomous inland water monitoring: Design and application of a surface vessel," *IEEE Robot. Autom. Magazine*, vol. 19, no. 1, pp. 62–72, 2012.

[20] http://github.com/jmetzen/bayesian_optimization (accessed 19/05/2016).

[21] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Adv. Neur. Inf. Proc. Sys.*, 2012.

[22] K. Swersky, J. Snoek, and R. P. Adams, "Freeze-thaw Bayesian optimization," 2014, http://arxiv.org/abs/1406.3896.

[23] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, "Lipschitzian optimization without the Lipschitz constant," *J. Opt. Theory Apps.*, vol. 79, no. 1, pp. 157–181, 1993.

[24] A. Krause and C. Guestrin, "A note on the budgeted maximization of submodular functions," 2005, Technical Report.

Supplementary Material

Truncated Variance Reduction: A Unified Approach to Bayesian Optimization and Level-Set Estimation (Ilija Bogunovic, Jonathan Scarlett, Andreas Krause, and Volkan Cevher, NIPS 2016)

A   Variations of the TRUVAR Algorithm

Our algorithm TRUVAR can naturally be adapted to suit various settings, including the following:

- **Non-monotonic M_t:** We have defined our sets M_t to become smaller on every time step. However, if $\beta_{(i)}$ is chosen aggressively, it may be preferable to replace M_{t-1} by D in (5)–(6), in which case some removed points may be added back in depending on how the posterior mean changes between steps. We take this approach in the real-world BO example of Section 4 in which the kernel hyperparameters are learned online, so as to avoid incorrectly ruling out points early on due to mismatched hyperparameters.

- **Avoiding computing the acquisition function everywhere:** We found that instead of computing the acquisition function at every point in D, limiting the selection to points in M_{t-1} has minimal effect on the performance. To reduce the computation even further, one could adopt a strategy such as that proposed in [22]: Take some relatively small number of points having the top GP-UCB or EI score, and then choose the point *in that restricted subset* having the highest score according to (3). In fact, the numerical results in Figure 3 suggest that this may not only reduce the computation, but also improve the performance in the very early rounds by making the algorithm *initially* behave more like GP-UCB or EI.

- **Pure variance reduction:** Setting $\eta_{(1)} = 0$ yields a *pure variance reduction* algorithm, which minimizes the total variance within M_t via a one-step lookahead. While our theory does not apply in this case, we found this choice to also work well in practice.

- **Implicit threshold for level-set estimation:** While we have focused on a threshold h for level-set estimation that is fixed in advance, one can easily incorporate the ideas of [2] to allow for an *implicit* threshold which is equal to some constant multiple of the function's maximum, which is random and unknown in advance.

- **Anticipating changes in M_t:** The acquisition function (3) computes the truncated variance reduction resulting from a one-step lookahead, but still sums over the previous set M_{t-1}. In order to make it more preferable to choose points that shrink M_t faster, it may be preferable to instead sum over $M_{t-1|x}$, defined to be the updated set upon adding x. The problem here is that such an update depend on the next posterior mean, whose update requires sampling f. One solution is to average over the measurement as in [8]; alternatively, a simpler approach is to replace the observed value with its mean when doing this one-step lookahead, and then using the true observed function sample only when x_t is actually chosen.

- **Continuous domains:** Our algorithm also extends to compact domains such as $[0,1]^d$. The main challenge is that the summations in (3) become integrals that need to be approximated numerically. The simplest way of doing this is to approximate the integrals by summations over a finite number of *representer points*, e.g., a grid of values that cover the domain sufficiently densely. The theoretical analysis of this modified algorithm is left for future work.

- **Batch setting:** As we show in the proof of Theorem 3.1, our algorithm can be interpreted as performing the first step of a greedy submodular covering problem at each time step. This leads to a very natural extension to the batch setting, in which multiple points (say, k of them) are chosen at each time step: Simply perform the first k steps of the greedy covering algorithm during each batch.

B   Further Details of Numerical Experiments

Other algorithms considered: We outline the algorithms that TRUVAR is compared against; full details can be found in the cited papers. For level-set estimation, we have the following:
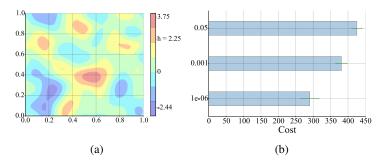
Figure 4: (a) Function used in synthetic level-set estimation experiments; (b) The total cost used by TRUVAR for each of the three noise levels.

- The GCHK algorithm [2] evaluates, at each iteration, the point that is not yet classified with the largest ambiguity: $x_t = \arg\max_{x \in M_{t-1}} \min\{u_t(x) - h, h - \ell_t(x)\}$, where u_t and ℓ_t are defined as in (4) with a parameter β_t replacing $\beta_{(i)}$. Here, similarly to our algorithm, M_{t-1} is the set of points that have not yet been classified as having a value above or below the threshold h.

- The straddle (STR) heuristic [7] chooses $x_t = \arg\max_{x \in D} 1.96\sigma_{t-1}(x) - |\mu_{t-1}(x) - h|$, favoring high-uncertainty points that are expected to have function values closer to h.

- The maximum variance rule (VAR) [2] simply chooses $x_t = \arg\max_{x \in D} \sigma_{t-1}(x)$.

For Bayesian optimization, we have the following:

- The expected improvement (EI) algorithm [1] chooses $x_t = \arg\max_{x \in D} \mathbb{E}_t[(f(x) - \xi_t)\mathbb{1}\{f(x) > \xi_t\}]$, where $\mathbb{E}_t[\cdot]$ denotes averaging with respect to the posterior distribution at time t, and ξ_t is the best observed value so far. Since the posterior is Gaussian, the expectation can easily be expressed in closed form.

- The Gaussian Process Upper Confidence Bound (GP-UCB) algorithm [3] chooses the points with the highest upper confidence bounds, $x_t = \arg\max_{x \in D} \mu_{t-1}(x) + \beta_t\sigma_{t-1}(x)$, where β_t is a parameter controlling the level of exploration performed.

- The Entropy Search (ES) algorithm can be interpreted as approximating the rule $x_t = \arg\min_{x \in D} h(f_{x^*|t,x})$, where $h(f) = \int_D f(x)\log\frac{1}{f(x)}dx$ denotes the differential entropy, and $f_{x^*|t,x}$ denotes the density function of the optimal action x^* given the observations up to time t along with the additional observation x. Intuitively, this rule seeks to minimize the uncertainty of x^*. Since its exact evaluation is intractable, it is approximated using Monte Carlo techniques to average with respect to the posterior distribution and the measurements.

- The Minimum Regret Search (MRS) algorithm [11] also resembles ES, but works with the expected regret instead of the differential entropy. Once again, Monte Carlo techniques are used to average with respect to the posterior distribution and the measurements.

Efficiently computing the acquisition function: To compute the value of the acquisition function (3) for different $x \in D$, we need to compute $\sigma_{t-1|x}^2(M_{t-1}) \in \mathbb{R}^{|M_{t-1}|}$, i.e., the posterior variance of points in M_{t-1} upon observing x along with $x_1, \cdots, x_{t-1}$. Instead of computing it directly, it is more efficient to recursively compute $\sigma_{t-1|x}^2(M_{t-1}) = \sigma_{t-1}^2(M_{t-1}) - \Delta_{t-1|x}(M_{t-1})$. The difference term, $\Delta_{t-1|x}(M_{t-1})$, can be computed as [8]:

$$\Delta_{t-1|x}(M_{t-1}) = \text{diag}\big(\text{Cov}_{t-1}(M_{t-1}, x)(\sigma^2 + \sigma_{t-1}^2(x))^{-1}\text{Cov}_{t-1}(M_{t-1}, x)^T\big), \quad (18)$$

where

$$\sigma_{t-1}^2(x) = k(x, x) - \mathbf{k}_{t-1}(x)^T\big(\mathbf{K}_{t-1} + \mathbf{\Sigma}_{t-1}\big)^{-1}\mathbf{k}_{t-1}(x) \quad (19)$$

$$\text{Cov}_{t-1}(M_{t-1}, x) = \mathbf{k}(M_{t-1}, x) - \mathbf{k}_{t-1}(M_{t-1})^T\big(\mathbf{K}_{t-1} + \mathbf{\Sigma}_{t-1}\big)^{-1}\mathbf{k}_{t-1}(x), \quad (20)$$

and where $\mathbf{k}(M_{t-1}, x) = [k(\overline{x}, x)]_{\overline{x} \in |M_{t-1}|} \in \mathbb{R}^{|M_{t-1}|}$, and $\mathbf{k}_{t-1}(M_{t-1}) = [k(\overline{x}, x)]_{\overline{x} \in |M_{t-1}|, x \in \{1,...,t-1\}} \in \mathbb{R}^{|M_{t-1}| \times (t-1)}$. When the Cholesky decomposition of $\mathbf{K}_{t-1} + \mathbf{\Sigma}_{t-1}$ is known, $\big(\mathbf{K}_{t-1} + \mathbf{\Sigma}_{t-1}\big)^{-1}\mathbf{k}_{t-1}(x)$ can be computed in time $O(t^2)$.

Details on LSE experiment with multiple noise levels: Figure 4a plots the randomly-generated function that was used in this experiment. Figure 4b plots the average cost spent by TRUVAR on each noise level by the end of the experiment, again averaged over 100 trials. We see that the cost is roughly equally distributed across the three levels. To be more specific, we observed that TRUVAR initially chooses high noise levels in order to cheaply explore, and throughout the course of the experiments, it gradually switches to lower noise levels in order to accurately determine the function values around the maximum. This is consistent with the behavior of the three version of GCHK, with $\sigma^2 = 0.05$ performing well in the early stages, but $\sigma^2 = 10^{-6}$ being preferable in the later stages.

Extension of TRUVAR for synthetic BO experiment We used the extension of TRUVAR to continuous domains outlined in Appendix A, approximating the integrals over M_t by summations that are restricted to points on a uniformly-spaced 50×50 grid covering $[0, 1]^2$. We optimized our acquisition function using DIRECT [23].

C Proof of General Result (Theorem 3.1)

We begin with the following lemma from [3].

Lemma C.1. [3] *For each t, define $\beta_t = 2 \log \frac{|D| t^2 \pi^2}{6\delta}$. With probability at least $1 - \delta$, we have for all x and t that $|f(x) - \mu_t(x)| \leq \beta_t^{1/2} \sigma_t(x)$.*

We conclude that in order for $\mu_t(\cdot) \pm \beta_{(i)}^{1/2} \sigma_t(\cdot)$ to provide valid confidence bounds, it suffices to ensure that $\beta_{(i)} \geq \beta_t$ for all t in epoch i. From (14), we see that this is true provided that the time taken to reach the end of the i-th epoch is at most $\frac{1}{c_{\min}} \sum_{i' \leq i} C_{(i')}$. Since $c_{\min}$ is the minimum pointwise cost, this holds provided that the cost incurred in epoch i is at most $C_{(i)}$. The bulk of the proof is devoted to showing that this is the case.

We connect TRUVAR with the following budgeted submodular covering problem:[2]

$$\text{minimize}_S \ c(S) \quad \text{subject to } g_t(S) = g_{t,\max}, \tag{21}$$

where

$$g_t(S) = \sum_{\overline{x} \in M_{t-1}} \max\left\{ \sigma_{t-1}^2(\overline{x}), \frac{\eta_{(i)}^2}{\beta_{(i)}} \right\} - \sum_{\overline{x} \in M_{t-1}} \max\left\{ \sigma_{t-1|S}^2(\overline{x}), \frac{\eta_{(i)}^2}{\beta_{(i)}} \right\}, \tag{22}$$

and where $g_{t,\max}$ is the highest possible value of $g_t(S)$ over arbitrarily large S, i.e., it is the value obtained once all of the summands in the second summation in (22) have saturated to $\frac{\eta_{(i)}^2}{\beta_{(i)}}$:

$$g_{t,\max} = \sum_{\overline{x} \in M_{t-1}} \left(\max\left\{ \sigma_{t-1}^2(\overline{x}), \frac{\eta_{(i)}^2}{\beta_{(i)}} \right\} - \frac{\eta_{(i)}^2}{\beta_{(i)}} \right) \tag{23}$$

$$= \sum_{\overline{x} \in M_{t-1}} \max\left\{ 0, \sigma_{t-1}^2(\overline{x}) - \frac{\eta_{(i)}^2}{\beta_{(i)}} \right\}. \tag{24}$$

We thus refer to $g_{t,\max}$ as the *excess variance*; see Figure 5 for an illustration. Note that each time instant t corresponds to a different function $g_t(S)$, and we are considering sets S of an arbitrary size even though our algorithm only chooses one point at each time instant.

By our assumption on the submodularity of the variance reduction function, and the fact that taking the minimum with a constant[3] preserves submodularity [17], $g_t(S)$ is also submodular. It is also easily seen to be monotonically increasing, and normalized in the sense that $g_t(\emptyset) = 0$.

---

[2]Recall that S may contain duplicates, and these are counted multiple times accordingly in the definitions of both $c(S)$ and $g_t(S)$. All of our equations can be cast in terms of standard sets (without duplicates) by expanding D to $D \times \{1, \cdots, N\}$ for any integer N that is larger than the maximum number of points that are chosen throughout the course of the algorithm.

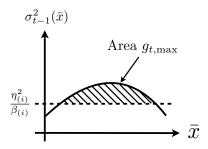[3]The minimum becomes a maximum after negation.

Figure 5: Illustration of the excess variance $g_{t,\max}$.

Our selection rule (3) at time t can now be interpreted as the first step in a greedy algorithm for solving the budgeted submodular optimization problem (21); specifically, the greedy rule optimizes the objective value per unit cost. To obtain performance guarantees, we use Lemma 2 of [24] specialized to $|S| = 1$, which reads as follows in our own notation:

$$g_t(\{x_t\}) \geq \frac{c(x_t)}{c(S_t^*)} g_{t,\max},\tag{25}$$

where S_t^* is an optimal solution to (21), and hence $g_t(S^*) = g_{t,\max}$. Here x_t is the point chosen greedily by our algorithm.

We now consider the behavior of the excess variance $g_{t,\max}$ in a single epoch, i.e., the duration of a single value of i in the algorithm. We claim that for t and $t+1$ in the same epoch, we have

$$g_{t+1,\max} \leq g_{t,\max} - g_t(\{x_t\}).\tag{26}$$

To see this, we note from (22)–(23) that this would hold with equality if we were to have $M_t = M_{t-1}$, since by definition we have $\sigma_t^2(\overline{x}) = \sigma_{t-1|\{x_t\}}^2(\overline{x})$. We therefore obtain (26) by recalling that M_t is decreasing in t with respect to inclusion, and noting from (24) that any given $g_{t,\max}$ can only decrease when M_{t-1} is smaller.

Combining (25)–(26) gives

$$g_{t+1,\max} \leq \left(1 - \frac{c(x_t)}{c(S_t^*)}\right) g_{t,\max}.\tag{27}$$

We also note that $c(S_{t+1}^*) \leq c(S_t^*)$, which follows since $\sigma_t^2(\cdot)$ is decreasing in t and M_t is shrinking in t, and therefore at time $t+1$ a smaller cost is required to ensure that all terms in the second summation of (22) have saturated to $\frac{\eta_{(i)}^2}{\beta_{(i)}}$. Hence, and applying (27) recursively, we obtain for t and $t+\ell$ in the same epoch that

$$\frac{g_{t+\ell,\max}}{g_{t,\max}} \leq \prod_{t'=t+1}^{t+\ell} \left(1 - \frac{c(x_{t'})}{c(S_t^*)}\right)\tag{28}$$

$$\leq \exp\left(-\frac{\sum_{t'=t+1}^{t+\ell} c(x_{t'})}{c(S_t^*)}\right),\tag{29}$$

where we have applied the inequality $1 - \alpha \leq e^{-\alpha}$. Moreover, the total cost incurred by choosing these points is precisely $\sum_{t'=t+1}^{t+\ell} c(x_t)$. Thus, letting $t_{(i)}$ be the first time index in the i-th epoch, we find that in order to remove all but a proportion γ of the initial excess variance $g_{t_{(i)},\max}$, it suffices that the cost incurred is at least

$$c(S_{t_{(i)}}^*) \log \frac{1}{\gamma}.\tag{30}$$

Next, we observe that since the posterior variance is upper bounded by one due to the assumption $k(x,x) = 1$, the initial excess variance $g_{t_{(i)},\max}$ is upper bounded by $g_{t_{(i)},\max} \leq |M_{t_{(i)}-1}|$, the size

of the set of potential maximizers at the *start* of the epoch. It follows that if we set

$$\gamma = \frac{\overline{\delta}^2 \eta_{(i)}^2}{\beta_{(i)} |M_{t_{(i)}-1}|}, \tag{31}$$

then removing all but a proportion γ of $g_{t_{(i)},\max}$ also implies removing all but $\overline{\delta}^2 \frac{\eta_{(i)}^2}{\beta_{(i)}}$ of it. In other words, if at time t we have incurred a cost in epoch i satisfying (30) with γ as in (31), then we must have $g_{t,\max} \le \overline{\delta}^2 \frac{\eta_{(i)}^2}{\beta_{(i)}}$.

Removing all of the excess variance would imply $\eta_{(i)}$-confidence at all points in M_t. In the worst case, the remaining excess variance $\overline{\delta}^2 \frac{\eta_{(i)}^2}{\beta_{(i)}}$ is concentrated entirely on a single point, in which case its confidence is upper bounded by $\sqrt{1 + \overline{\delta}^2} \eta_{(i)}$, which is further upper bounded by $(1 + \overline{\delta})\eta_{(i)}$ due to the identity $\sqrt{1 + \alpha^2} \le 1 + \alpha$.

Combining these observations, we conclude that in the i-th epoch, upon incurring a cost of at least

$$c(S_{t_{(i)}}^*) \log \frac{|M_{t_{(i)}-1}|\beta_{(i)}}{\overline{\delta}^2 \eta_{(i)}^2}, \tag{32}$$

we are guaranteed to have $(1 + \overline{\delta})\eta_{(i)}$-confidence for all points in M_t. Having such confidence is precisely the condition used in the algorithm to move onto the next epoch, and we conclude that the epoch must end by (or sooner than) the time that (32) holds.

In accordance with the discussion following Lemma C.1, we need to show that when the high-probability event in that lemma holds true, (32) is upper bounded by the right-hand side of (13) for all epochs. We do this via an induction argument on the epoch number:

- As a base case, recalling that $M_0 = M_{(0)} = D$, we find that (32) and (13) coincide, with the addition of $c_{\max}$ arising since once (32) is exceeded, it may be exceeded by any amount up to $c_{\max}$.

- Fix an epoch number $i > 1$, and suppose that for all $i' < i$, the cost incurred in epoch i' was at most $C^*\big(\frac{\eta_{(i')}}{\beta_{(i')}^{1/2}}, \overline{M}_{(i'-1)}\big) \log \frac{|\overline{M}_{(i'-1)}|\beta_{(i')}}{\overline{\delta}^2 \eta_{(i')}^2} + c_{\max}$. By the choice of $\beta_{(i')}$ in (14), we find that under the event in Lemma C.1, the confidence bounds $\mu_t \pm \sqrt{\beta_{(i')}}\sigma_t$ must have been valid for all t in the epochs $i' < i$, and hence $\overline{M}_{(i-1)} \subseteq M_{t_{(i)}-1}$ (*cf.*, (9)–(10)).

  From this, we claim that an analogous argument to the base case implies that (32) is upper bounded by the right-hand side of (13), as required. The only additional argument compared to the base case is noting that $c(S_{t_{(i)}}^*)$ defines the minimum cost to uniformly shrink the posterior standard deviation within $M_{t_{(i)}-1}$ down to $\frac{\eta_{(i)}^2}{\beta_{(i)}}$ after already having chosen $x_1, \ldots, x_{t_{(i)}-1}$, whereas $C^*\big(\frac{\eta_{(i)}}{\beta_{(i)}^{1/2}}, \overline{M}_{(i-1)}\big)$ is defined analogously for the set $\overline{M}_{(i-1)}$ with no previously-chosen points. The latter clearly upper bounds the former.

Finally, we check the conditions for ϵ-accuracy in Definition 3.1. In the case of BO, summing (13) over all of the epochs such that $4(1 + \overline{\delta})\eta_{(i-1)} > \epsilon$ yields (15); recall from (9) that after any epoch i such that $4(1 + \overline{\delta})\eta_{(i)} \le \epsilon$, all points are at most ϵ-suboptimal. We also note that all true maxima must remain in M_t, due to the fact that we showed $\beta_{(i)}$ yields valid confidence bounds with high probability, and we only ever discard points that are deemed suboptimal according to those bounds. For LSE, a similar conclusion follows from (10) by summing over all epochs such that $2(1 + \overline{\delta})\eta_{(i-1)} > \frac{\epsilon}{2}$, which is the same as $4(1 + \overline{\delta})\eta_{(i-1)} > \epsilon$. Once again, all points in H_t and L_t are correct due to the validity of our confidence bounds.

14

D Simplified Result for the Homoscedastic and Unit-Cost Setting

Since we are focusing on unit costs $c(x) = 1$, the cost simply corresponds to the number of rounds T. To highlight this fact, we replace C^* in (11) by

$$T^*(\xi, M) = \min_S \left\{ |S| \; : \; \max_{\overline{x} \in M} \sigma_{0|S}(\overline{x}) \leq \xi \right\}, \tag{33}$$

and similarly replace (34)–(36) by

$$T_{(i)} \geq T^*\left(\frac{\eta_{(i)}}{\beta_{(i)}^{1/2}}, \overline{M}_{(i-1)} \right) \log \frac{|\overline{M}_{(i-1)}| \beta_{(i)}}{\overline{\delta}^2 \eta_{(i)}^2} + 1 \tag{34}$$

$$\beta_{(i)} \geq 2 \log \frac{|\overline{M}_{(i-1)}| \left(\sum_{i' \leq i} T_{(i')} \right)^2 \pi^2}{6\delta} \tag{35}$$

$$T_\epsilon = \sum_{i \,:\, 4(1+\overline{\delta})\eta_{(i-1)} > \epsilon} T_{(i)}. \tag{36}$$

In this section, we prove the following as an application of Theorem 3.1.

Corollary D.1. *Fix $\epsilon > 0$ and $\delta \in (0, 1)$, define $\beta_T = 2 \log \frac{|D| T^2 \pi^2}{6\delta}$, and set $\eta_{(1)} = 1$ and $r = \frac{1}{2}$. There exist choices of $\beta_{(i)}$ (not depending on the time horizon T) such that we have ϵ-accuracy with probability at least $1 - \delta$ once the following condition holds:*

$$T \geq \left(C_1 \gamma_T \beta_T \frac{96(1+\overline{\delta})^2}{\epsilon^2} + 2 \left\lceil \log_2 \frac{8(1+\overline{\delta})}{\epsilon} \right\rceil \right) \log \frac{16(1+\overline{\delta})^2 |D| \beta_T}{\overline{\delta}^2 \epsilon^2}, \tag{37}$$

where $C_1 = \frac{1}{\log(1+\sigma^{-2})}$. This condition is of the form $T \geq \Omega^*\left(\frac{C_1 \gamma_T \beta_T}{\epsilon^2} + 1 \right)$.

We bound the cardinality of S in (33) by considering a procedure that greedily picks $\arg\max_{x \in M} \sigma_t(x)$. We claim that after selecting k points according to this procedure to construct a set S_k, we have

$$\max_x \sigma_{0|S_k}^2(x) \leq C_1 \frac{\gamma_k}{k}, \tag{38}$$

where $C_1 = \frac{1}{\log(1+\sigma^{-2})}$. This is seen by writing

$$k \max_{x \in M} \sigma_{0|S_k}^2(x) = k \sigma_{0|S_k}^2(x_k) \tag{39}$$

$$\leq \sum_{j=1}^k \sigma_{0|S_j}^2(x_j) \tag{40}$$

$$\leq \frac{1}{\log(1 + \sigma^{-2})} \gamma_k, \tag{41}$$

where we respectively used that x_k maximizes $\sigma_{0|S_k}$, that $\sigma_{0|S_i}(x_i)$ always decreases as more points are chosen, and the bound on the sum of variances of sampled points from [3, Lemma 5.4].

Identifying k with T^*, and $\max_{x \in M} \sigma_{0|S_k}(x)$ with $\xi = \frac{\eta}{\beta^{1/2}}$ (for some η and β to be specified), we obtain from (41) that

$$T^*\left(\frac{\eta}{\beta^{1/2}}, M \right) \leq \min \left\{ T^* \; : \; T^* \geq \frac{C_1 \gamma_{T^*} \beta}{\eta^2} \right\}. \tag{42}$$

Consider the value $T^*\left(\frac{\eta_{(i)}}{\beta_{(i)}^{1/2}}, \overline{M}_{(i-1)} \right)$ corresponding to the parameters $\eta = \eta_{(i)}$ and $\beta = \beta_{(i)}$ associated with epoch i. Letting $T = T_\epsilon$ denote the total time horizon, and using (36), we find that $\beta_{(i)}$ in (35) can be upper bounded by $2 \log \frac{|D| T^2 \pi^2}{6\delta}$, which is precisely β_T. By similarly using the monotonicity of γ_t, we obtain

$$T^*\left(\frac{\eta_{(i)}}{\beta_{(i)}^{1/2}}, \overline{M}_{(i-1)} \right) \leq \frac{C_1 \gamma_T \beta_T}{\eta_{(i)}^2} + 1, \tag{43}$$

where the addition of one is to account for possible rounding up to the nearest integer.

Using (43), we find that in order for (34) to hold it suffices that

$$T_{(i)} \geq \left(\frac{C_1 \gamma_T \beta_T}{\eta_{(i)}^2} + 1 \right) \log \frac{|\overline{M}_{(i-1)}|\beta_{(i)}}{\overline{\delta}^2 \eta_{(i)}^2} + 1. \tag{44}$$

Since we are only considering values of i such that $4(1+\overline{\delta})\eta_{(i-1)} > \epsilon$, and since $\overline{M}_{(i-1)} \subseteq D$, we can upper bound the logarithm by $\log \frac{16(1+\overline{\delta})^2|D|\beta_{(i)}}{\overline{\delta}^2 \epsilon^2} > 1$, and hence in order for (44) to hold it suffices that

$$T_{(i)} \geq \left(\frac{C_1 \gamma_T \beta_T}{\eta_{(i)}^2} + 2 \right) \log \frac{16(1+\overline{\delta})^2|D|\beta_T}{\overline{\delta}^2 \epsilon^2}. \tag{45}$$

We also note that since $\eta_{(i)} = \eta_{(1)} r^{i-1}$, the condition $4(1+\overline{\delta})\eta_{(i-1)} > \epsilon$ is equivalent to

$$4(1+\overline{\delta})\eta_{(1)} r^{i-2} > \epsilon \tag{46}$$

$$\iff r^{i-2} > \frac{\epsilon}{4(1+\overline{\delta})\eta_{(1)}} \tag{47}$$

$$\iff i < 2 + \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \tag{48}$$

$$\iff i \leq \left\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \right\rceil + 1 \tag{49}$$

$$\iff i \leq \left\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{r\epsilon} \right\rceil, \tag{50}$$

where in the last line we used $\log_{1/r} \frac{1}{r} = 1$. Summing (45) over all such i in accordance with (36), we obtain following sufficient condition on the time horizon for ϵ-accuracy:

$$T \geq \left(C_1 \gamma_T \beta_T \sum_{i=1}^{\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \rceil + 1} \frac{1}{\eta_{(i)}^2} + 2 \left\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{r\epsilon} \right\rceil \right) \log \frac{16(1+\overline{\delta})^2|D|\beta_T}{\overline{\delta}^2 \epsilon^2}. \tag{51}$$

Finally, we weaken this condition by upper bounding the summation as follows:

$$\sum_{i=1}^{\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \rceil + 1} \frac{1}{\eta_{(i)}^2} = \sum_{i=1}^{\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \rceil + 1} \frac{1}{\eta_{(1)}^2 r^{2(i-1)}} \tag{52}$$

$$= \sum_{i=0}^{\lceil \log_{1/r} \frac{4(1+\overline{\delta})\eta_{(1)}}{\epsilon} \rceil} \frac{1}{\eta_{(1)}^2 r^{2i}} \tag{53}$$

$$\leq \frac{1}{r^2(1-r^2)} \frac{16(1+\overline{\delta})^2}{\epsilon^2}, \tag{54}$$

where the last line follows from the identity $\sum_{i=0}^{\lceil \log_{1/r} A \rceil} \frac{1}{r^{2i}} \leq \frac{1}{r^2(1-r^2)} A^2$ for $\log_{1/r} A > 0$. Substituting $r = \frac{1}{2}$ and $\eta_{(1)} = 1$ concludes the proof; the former yields $\frac{1}{r^2(1-r^2)} = \frac{16}{3} \leq 6$.

E   Proof of Improved Noise Dependence (Corollary 3.1))

The bound in (41) is based on the inequality [3, Lemma 5.4]

$$\frac{\sigma_t^2}{\sigma^2} \leq C_1 \log \left(1 + \frac{\sigma_t^2}{\sigma^2} \right) \tag{55}$$

for $\sigma_t^2 \in [0,1]$ (with $C_1 = \frac{\sigma^{-2}}{\log(1+\sigma^{-2})}$), which can be very loose when σ^2 is small. Our starting point to improve the noise dependence is to note that the following holds under the more restrictive

condition $\sigma_t^2 \leq \sigma^2$:

$$\sigma_t^2 = \sigma^2 \frac{\sigma_t^2}{\sigma^2} \tag{56}$$

$$\leq 2\sigma^2 \log\left(1 + \frac{\sigma_t^2}{\sigma^2}\right), \tag{57}$$

where we have used the fact that $\alpha \leq 2\log(1+\alpha)$ for $\alpha \in [0,1]$.

The idea now is to use (57) in the epochs that are late enough so that $\sigma_t^2 \leq \sigma^2$, and (41) in the earlier epochs. Since $(1 + \bar{\delta})\eta_{(i)}$ is the confidence level obtained after epoch i, and since $\beta_{(i)}^{1/2}\sigma_t$ is the confidence level after time t, we find that in order to ensure $\sigma_t^2 \leq \sigma^2$ it suffices that $\frac{(1+\bar{\delta})^2 \eta_{(i)}^2}{\beta_{(i)}} \leq \sigma^2$. Moreover, our choice of $\beta_{(i)}$ in (14) is always greater than one when $|D| \geq 2$ (which is a trivial condition), and hence we can weaken this condition to $(1 + \bar{\delta})^2 \eta_{(i)}^2 \leq \sigma^2$, and write

$$\sum_i T_{(i)} \leq \sum_{i\,:\,(1+\bar{\delta})^2 \eta_{(i-1)}^2 > \sigma^2} T_{(i)}^{(C_1)} + \sum_i T_{(i)}^{(2\sigma^2)}, \tag{58}$$

where $T_{(i)}^{(C_1)}$ denotes bound on $T_{(i)}$ in (45) based on (41), and $T_{(i)}^{(C_1)}$ denotes the analogous bound based on (57) with $2\sigma^2$ in place of C_1. Similarly to (50), the first summation is over a subset of the range $i \leq \lceil \log_{1/r} \frac{(1+\bar{\delta})\eta_{(1)}}{r\sigma} \rceil$, and it follows that the condition (51) may be replaced by

$$T \geq \left(2\sigma^2 \gamma_T \beta_T \sum_{i=1}^{\lceil \log_{1/r} \frac{4(1+\bar{\delta})\eta_{(1)}}{\epsilon}\rceil+1} \frac{1}{\eta_{(i)}^2} + 2\left\lceil \log_{1/r} \frac{4(1+\bar{\delta})\eta_{(1)}}{r\epsilon}\right\rceil\right) \log \frac{16(1+\bar{\delta})^2|D|\beta_T}{\bar{\delta}^2 \epsilon^2}$$

$$+ \left(C_1 \gamma_T \beta_T \sum_{i=1}^{\lceil \log_{1/r} \frac{(1+\bar{\delta})\eta_{(1)}}{\sigma}\rceil+1} + 2\left\lceil \log_{1/r} \frac{(1+\bar{\delta})\eta_{(1)}}{r\sigma}\right\rceil\right) \log \frac{16(1+\bar{\delta})^2|D|\beta_T}{\bar{\delta}^2 \epsilon^2}. \tag{59}$$

The first summation is handled in the same way as the previous subsection, and the second summation is upper bounded by writing

$$\sum_{i=1}^{\lceil \log_{1/r} \frac{(1+\bar{\delta})\eta_{(1)}}{\sigma}\rceil+1} \frac{1}{\eta_{(i)}^2} = \sum_{i=1}^{\lceil \log_{1/r} \frac{(1+\bar{\delta})\eta_{(1)}}{\sigma}\rceil+1} \frac{1}{\eta_{(1)}^2 r^{2(i-1)}} \tag{60}$$

$$\leq \frac{(1+\bar{\delta})^2}{r^2(1-r^2)} \frac{1}{\sigma^2}, \tag{61}$$

where (60) follows since $\eta_{(i)} = \eta_{(1)} r^{i-1}$, and (61) follows in the same way as (54). Once again, setting $r = \frac{1}{2}$ and $\eta_{(1)} = 1$ concludes the proof, with the third term in (17) coming from the identity $2\lceil \log_2 \frac{8(1+\bar{\delta})}{\epsilon}\rceil + 2\lceil \log_2 \frac{2(1+\bar{\delta})}{\sigma}\rceil \leq 2\lceil \log_2 \frac{32(1+\bar{\delta})^2}{\epsilon\sigma}\rceil$.

F Proof for the Setting of Choosing Noise (Corollary 3.2)

The proof follows the same arguments as those of Appendices D and E, with C^* being upper bounded in K different ways, one for each possible noise level. The choice $\beta_T = 2\log \frac{|D|T^2 c_{\max}^2 \pi^2}{6\delta c_{\min}^2}$ arises as a simple upper bound to the right-hand side of (14) resulting from the fact that $\sum_{t=1}^T c(x_t) \leq c_{\max}T$.