Supplementary Material - Exploiting Strategy-Space Diversity for Batch Bayesian Optimization

Sunil Gupta **Alistair Shilton** **Santu Rana** **Svetha Venkatesh**
{sunil.gupta, alistair.shilton, santu.rana, svetha.venkatesh}@deakin.edu.au
Deakin University, Geelong, Australia

1 Preliminaries

At every iteration of our proposed batched BO algorithm we solve ($\mathcal{X}_t^{\rm p}$ in the Pareto sense):

$$\begin{array}{rl} x_t^{\rm u} &= \mathrm{argmax}_{x\in\mathbb{X}} \left(\mu_{t-1}\left(x\right)+\sqrt{{}_2\beta_t}\sigma_{t-1}\left(x\right)\right) \\ \mathcal{X}_t^{\rm p} &\subseteq \mathrm{argmax}_{x\in\mathbb{X}_t^+}\left(\mu_{t-1}\left(x\right),\sigma_{t-1}\left(x\right)\right) \end{array} \tag{1}$$

to form $\mathcal{X}_t = \left\{x_t^1, x_t^2, \ldots, x_t^{n_t}\right\} \subset \mathcal{X}_t^{\rm p} \cup \{x_t^{\rm u}\}$, $x_t^{\rm u} \in \mathcal{X}_t$, as per algorithm 2. We use the shorthand $\mu_t^i = \mu_{t-1}(x_t^i)$, $\sigma_t^i = \sigma_{t-1}(x_t^i)$ etc.

It follows from the properties of Pareto optimality that each x_t^i is the maximum of a scalarised acquisition function a_t^i characterised by β_t^i:

$$x_t^i = \mathrm{argmax}_{x\in\mathbb{X}}\left(a_t^i\left(x\right)=\mu_t\left(x\right)+\sqrt{\beta_t^i}\sigma_t\left(x\right)\right)$$

We assume without loss of generality that:

$$\sigma_t^1 \leq \sigma_t^2 \leq \ldots \leq \sigma_t^{n_t} \;\forall t \tag{2}$$

from which we obtain the following results:

Lemma 1 *For all t, $j \neq k \leq n_t$:*

$$\begin{array}{l} \left(\beta_t^j < \beta_t^k\right) \Rightarrow \left(\sigma_t^j \leq \sigma_t^k\right) \\ \left(\sigma_t^j < \sigma_t^k\right) \Rightarrow \left(\beta_t^j \leq \beta_t^k\right) \end{array}$$

Proof: We first prove that $(\beta_t^j < \beta_t^k) \Rightarrow (\sigma_t^j \leq \sigma_t^k)$. By definition $\mu_t^i + \sqrt{\beta_t^i}\sigma_t^i$ is the maximum of the acquisition function a_t^i $\forall i$. So:

$$\begin{array}{rl} \mu_t^j + \sqrt{\beta_t^j}\sigma_t^j &\geq \mu_t^k + \sqrt{\beta_t^j}\sigma_t^k \\ \mu_t^j + \sqrt{\beta_t^k}\sigma_t^j &\leq \mu_t^k + \sqrt{\beta_t^k}\sigma_t^k \end{array} \tag{3}$$

Hence:

$$\begin{array}{rl} -\mu_t^j - \sqrt{\beta_t^j}\sigma_t^j &\leq -\mu_t^k - \sqrt{\beta_t^j}\sigma_t^k \\ \mu_t^j + \sqrt{\beta_t^k}\sigma_t^j &\leq \mu_t^k + \sqrt{\beta_t^k}\sigma_t^k \end{array}$$

so:

$$\left(\sqrt{\beta_t^k} - \sqrt{\beta_t^j}\right)\sigma_t^j \leq \left(\sqrt{\beta_t^k} - \sqrt{\beta_t^j}\right)\sigma_t^k$$

and it follows that $\sigma_t^j \leq \sigma_t^k$.

To demonstrate that $(\sigma_t^j < \sigma_t^k) \Rightarrow (\beta_t^j \leq \beta_t^k)$ suppose the converse - that is, $(\sigma_t^j < \sigma_t^k)$ but $(\beta_t^j > \beta_t^k)$. Let $r = \sigma_t^k - \sigma_t^j > 0$ and $s = \sqrt{\beta_t^j} - \sqrt{\beta_t^k} > 0$. Clearly (3) must still hold, so:

$$\begin{array}{rl} \mu_t^j + \sqrt{\beta_t^j}\sigma_t^j &\geq \mu_t^k + \sqrt{\beta_t^j}\sigma_t^j + \sqrt{\beta_t^j}r \\ \mu_t^j + \sqrt{\beta_t^j}\sigma_t^j &\leq \mu_t^k + \sqrt{\beta_t^j}\sigma_t^j + \sqrt{\beta_t^j}r - sr \end{array}$$

Hence:

$$\begin{array}{rl} -\mu_t^j - \sqrt{\beta_t^j}\sigma_t^j &\leq -\mu_t^k - \sqrt{\beta_t^j}\sigma_t^j - \sqrt{\beta_t^j}r \\ \mu_t^j + \sqrt{\beta_t^j}\sigma_t^j &\leq \mu_t^k + \sqrt{\beta_t^j}\sigma_t^j + \sqrt{\beta_t^j}r - sr \end{array}$$

so $sr \leq 0$, which contradicts the definition of s, r. It follows by contradiction that $(\sigma_t^j < \sigma_t^k) \Rightarrow (\beta_t^j \leq \beta_t^k)$. $\square$

Lemma 2 *Following the definitions $\forall t$:*

$${}_2\beta_t = \beta_t^1 \leq \beta_t^2 \leq \ldots \leq \beta_t^{n_t}$$

Proof: This follows from assumption (2), theorem 1 and definition 1 (which implies that $\sigma_{t-1}(x_t^{\rm u}) < \sigma_{t-1}(x_t^{\rm p}))$ $\forall x_t^{\rm p} \in \mathcal{X}_t^{\rm p}$, and hence $\beta_t^i \geq {}_2\beta_t$ $\forall i$). $\square$

Corollary 1 *For all i the scalarisation constants β_t^i are generated by a relaxed GP-UCB strategy.*

Proof: Define $\gamma_t^i = \beta_t^i - {}_2\beta_t$ $\forall i$. We see from Lemma 2 that $\gamma_t^i \geq 0$. As ${}_2\beta_t$ is generated by the standard GP-UCB strategy, ${}_2\beta_t = 2\log(\eta_t {}_2\pi_t/\delta)$, where ${}_2\pi_t = \pi^2 t^2/6$ and $\sum_t 1/{}_2\pi_t = 1$. Hence:

$$\beta_t^i = 2\log\left(\tfrac{\eta_t}{\delta}{}_2\pi_t\right) + \gamma_t^i = 2\log\left(\tfrac{\eta_t}{\delta}\pi_t\right)$$

where $\pi_t = {}_2\pi_t \exp(\gamma_t^i/2)$. $\pi_t \geq {}_2\pi_t$ as $\gamma_t^i \geq 0$, so $\sum_t 1/\pi_t \leq 1$. Hence β_t^i is generated by a relaxed GP-UCB strategy $\forall i$. □

Lemma 3 *Following the definitions:*

$$\textstyle\sum_{t \leq T} \sigma_t^1 \leq \sum_{t \leq T} \frac{1}{n_t} \sum_{1 \leq i \leq n_t} \sigma_t^i$$

Proof: This follows from (2). □

2 Regret Bounds

As defined in the main manuscript the regret for a single evaluation $x_t^i \in \mathcal{X}_t$ is:

$$r_t^i = f(x^*) - f\left(x_t^i\right)$$

where x^* is the maximiser for f; and the simple regret for the complete batch $\mathcal{X}_t$ is:

$$r_t = \min_i r_t^i$$

and:

$$\begin{aligned} R_T^n &= \textstyle\sum_{t \leq T} r_t && \text{(batch cumulative regret)} \\ R_{Tn} &= \textstyle\sum_{t \leq T} \sum_{i \leq n_t} r_t^i && \text{(total cumulative regret)} \end{aligned}$$

are the batch cumulative regret and total cumulative regret, respectively.

We have the following results:

Lemma 4 *If $\mathbb{X}$ is finite; or if $\mathbb{X}$ is compact and convex and $k(x,x)) \leq 1$ $\forall x \in \mathbb{X}$; then for all t the bound:*

$$r_t \leq r_t^1 \leq 2\sqrt{{}_2\beta_t}\sigma_t^i + D_t$$

holds with probability $\geq 1 - \delta$, where $D_t = 0$ if $|\mathbb{X}|$ is finite, $D_t = \frac{1}{t^2}$ otherwise.

Proof: The proof is similar to that of Lemma 5.2, 5.4 and 5.8 in Srinivas et al. (2012) and Lemma 1 in Contal et al. (2013). We consider only the finite $|\mathbb{X}|$ case here (the infinite case follows by analogy with the method in Srinivas et al. (2012)). By definition $r_t \leq r_t^1$. As β_t^p is generated by a relaxed GP-UCB strategy we have that $|f(x) - \mu_{t-1}(x)| \leq \sqrt{\beta_t^p}\sigma_{t-1}(x)$ holds with probability $\geq 1 - \delta$ $\forall x, p$ (Lemma 5.1, Srinivas et al. (2012)). We also have from Lemma 2 that $\beta_t^1 = {}_2\beta_t$. It follows that, with probability $\geq 1 - \delta$:

$$\begin{aligned} &f(x^*) \leq \mu_{t-1}(x^*) + \sqrt{{}_2\beta_t}\sigma_{t-1}(x^*) \leq \mu_t^1 + \sqrt{{}_2\beta_t}\sigma_t^1 \\ &\mu_t^1 - \sqrt{{}_2\beta_t}\sigma_t^1 \leq f\left(x_t^1\right) \leq \mu_t^1 + \sqrt{{}_2\beta_t}\sigma_t^1 \end{aligned}$$

and hence:

$$\begin{aligned} r_t \leq r_t^1 &= f(x^*) - f\left(x_t^1\right) \\ &\leq \mu_t^1 + \sqrt{{}_2\beta_t}\sigma_t^1 - \mu_t^1 + \sqrt{{}_2\beta_t}\sigma_t^1 \\ &= 2\sqrt{{}_2\beta_t}\sigma_t^1 \end{aligned}$$

holds with probability $\geq 1 - \delta$. □

Lemma 5 *For all T the bound:*

$$\textstyle\sum_{t \leq T} \sum_{j \leq n_t} \left(\sigma_t^j\right)^2 \leq C_1' \gamma_{A_T T}$$

holds with probability $\geq 1 - \delta$, where $C_1' = 2/\log\left(1 + \nu^{-2}\right)$; $A_T = \frac{1}{T}\sum_{t \leq T} n_t$ is the arithmetic mean of the batch sizes; and $\gamma_{A_T T}$ is the max information gain obtainable from a sequence of length $A_T T$.

Proof: The proof follows that of Lemmas 5.3 and 5.4 in Srinivas et al. (2012), as per Contal et al. (2013), except that in this case the batch size varies with t, leading to $A_T T$ term in the result. □

Theorem 1 *Let $\delta \in]0, 1[$. Assuming $k(x,x) \leq 1$ $\forall x \in \mathbb{X}$ and either $\mathbb{X}$ is either finite or $\mathbb{X} \subset [0, r]^d$ and f satisfies $\Pr\left\{\sup_{x \in \mathbb{X}} |\partial f / \partial x_j| > L\right\} \leq a e^{-(L/b)^2}$ then for all T:*

$$R_T^n \leq \sqrt{\tfrac{T}{H_T} C_1 \, {}_2\beta_T \gamma_{A_T T}} + C_2 \tag{4}$$

holds with probability $\geq 1 - \delta$, where $A_T = \frac{1}{T}\sum_{t \leq T} n_t$ and $H_T = \frac{T}{\sum_{t \leq T} \frac{1}{n_t}}$ are, respectively, the arithmetic and harmonic means of the batch sizes; $C_1 = 8/\log\left(1 + \nu^{-2}\right)$; $C_2 = 0$ if $|\mathbb{X}|$ is finite, $\pi^2/6$ otherwise; and $\gamma_{A_T T}$ is the max information gain obtainable from a sequence of length $A_T T$.

Proof: Applying previous results, as per Contal et al. (2013):

$$\begin{aligned} R_T^n &= \textstyle\sum_{t \leq T} r_t \\ &\leq \textstyle\sum_{t \leq T} 2\sqrt{{}_2\beta_t}\sigma_t^1 + C_2 \text{ (Lem 4)} \\ &\leq 2\sqrt{{}_2\beta_T} \textstyle\sum_{t \leq T} \sigma_t^1 + C_2 \text{ (${}_2\beta_t$ increasing)} \\ &\leq 2\sqrt{{}_2\beta_T} \textstyle\sum_{t \leq T} \frac{1}{n_t} \sum_{i \leq n_t} \sigma_t^i + C_2 \text{ (Lem 3)} \\ &\leq 2\sqrt{{}_2\beta_T} \sqrt{\tfrac{T}{H_T} \textstyle\sum_{t \leq T} \sum_{i \leq n_t} \left(\sigma_t^i\right)^2} + C_2 \text{ (C.-S.)} \\ &\leq 2\sqrt{{}_2\beta_T} \sqrt{\tfrac{T}{H_T} C_1' \gamma_{A_T T}} + C_2 \text{ (Lem 5)} \\ &\leq \sqrt{\tfrac{T}{H_T} C_1 \, {}_2\beta_T \gamma_{A_T T}} + C_2 \end{aligned}$$

where $C_2 = \sum_{t \leq T} D_t$. Hence $C_2 = 0$ if $|\mathbb{X}|$ is finite, $C_2 = \zeta(2) = \pi^2/6$ otherwise. □

To this point our proof would apply if we replaced the relevant region $\mathbb{X}_t^+$ in our algorithm with $\mathbb{X}$. To finish our proof with respect to bounding total cumulative regret R_{Tn}, however, requires the use of the relevant region. We begin with an analogue to Lemma 4:

Lemma 6 *If $\mathbb{X}$ is finite; or if $\mathbb{X}$ is compact and convex and $k(x,x)) \leq 1$ $\forall x \in \mathbb{X}$; then for all t the bound:*

$$r_t \leq r_t^i \leq 6\sqrt{{}_2\beta_t}\sigma_t^i + D_t$$

holds with probability $\geq 1 - \delta$, where $D_t = 0$ if $|\mathbb{X}|$ is finite, $D_t = \frac{1}{t^2}$ otherwise.

Proof: This proof follows analogously by the method of Contal et al. (2013), proof of Lemma 6. □

This allows us to complete our proof:

Theorem 2 *Under the conditions of theorem 1:*

$$R_{Tn} \leq \sqrt{A_T T C_3 \, {}_2\beta_T \gamma_{A_T T}} + N_\infty C_2 \tag{5}$$

holds with probability $\geq 1-\delta$, *where* $A_T = \frac{1}{T}\sum_{t\leq T} n_t$ *is the arithmetic mean of the batch sizes;* $N_\infty = \max_{t\leq T} n_t$; $C_2 = 0$ *if* $|\mathbb{X}|$ *is finite,* $\pi^2/6$ *otherwise;* $C_3 = 72/\log(1+\nu^{-2})$*; and* $\gamma_{A_T T}$ *is the max information gain obtainable from a sequence of length* $A_T T$.

Proof: Applying previous results, as per Contal et al. (2013):

$$\begin{aligned} R_{Tn} &= \textstyle\sum_{t\leq T}\sum_{i\leq n_t} r_t^i \\ &\leq \textstyle\sum_{t\leq T}\sum_{i\leq n_t} 6\sqrt{{}_2\beta_t}\sigma_t^i + E \text{ (Lem 6)} \\ &\leq 6\sqrt{{}_2\beta_T}\textstyle\sum_{t\leq T}\sum_{i\leq n_t}\sigma_t^i + E \text{ (}{}_2\beta_t\text{ increasing)} \\ &\leq 6\sqrt{{}_2\beta_T}\sqrt{A_T T\textstyle\sum_t\sum_i\left(\sigma_t^i\right)^2} + E \text{ (C.-S.)} \\ &\leq 6\sqrt{{}_2\beta_T}\sqrt{A_T T C_1' \gamma_{A_T T}} + E \text{ (Lem 5)} \\ &\leq \sqrt{A_T T C_3 \, {}_2\beta_T \gamma_{A_T T}} + E \end{aligned}$$

where $E = \sum_{t\leq T}\sum_{i\leq n_t} D_t \leq N_\infty C_2$. □

3 A Note on Relevant Region Selection for (only) Batch Cumulative Regret Bounding

As noted in the previous section, the proof up to and including theorem 1 will hold if we use for our relevant region the (unconstrained) set $\mathbb{X}$ - that is:

$$\begin{aligned} &x_t^{\mathrm{u}} = \mathrm{argmax}_{x\in\mathbb{X}}\left(\mu_{t-1}(x) + \sqrt{{}_2\beta_t}\sigma_{t-1}(x)\right) \\ &\mathcal{X}_t^{\mathrm{p}} \subseteq \mathcal{X}_t^{\mathrm{pf}} \subseteq \mathrm{argmax}_{x\in\mathbb{X}}\left(\mu_{t-1}(x), \sigma_{t-1}(x)\right) \\ &\mathcal{X}_t = \mathcal{X}_t^{\mathrm{p}} \cup \{x_t^{\mathrm{u}}\} \end{aligned} \tag{6}$$

The caveat in this case is that Lemma 2 must be replaced by:

Lemma 7 *Following the definitions* $\forall t$:

$${}_2\beta_t \leq \beta_t^{m_t} \leq \beta_t^{m_t+1} \leq \ldots \leq \beta_t^{n_t}$$

where $1 \leq m_t \leq n_t$.

In subsequent results only those $n_t' = m_t - n_t + 1$ recommendations with the largest variances σ_t^i are included. The bound (4) on batch cumulative regret:

$$R_T^n \leq \sqrt{\tfrac{T}{H_T'} C_1 \, {}_2\beta_T \gamma_{A_T' T}} + C_2 \tag{7}$$

will hold in this case where $A_T' = \frac{1}{T}\sum_{t\leq T} n_t'$ and $H_T' = \frac{T}{\sum_{t\leq T}\frac{1}{n_t'}}$ are, respectively, the arithmetic and harmonic means of the number of recommendation in each batch *that are at least as exploratory as the standard GP-UCB recommendation* x_t^{u}.

Note that an analogous result does not follow for total cumulative regret R_{Tn}, as in this case the proof hinges on the regret for all points in the recommendation set being bounded, which cannot be guaranteed for recommendations not generated by a relaxed GP-UCB strategy (i.e. Lemma 6, and subsequently Theorem 2, does not hold in this case).

4 Additional Experimental Results

4.1 Evolution of Batch size

We provide additional detail into the batch size chosen by our proposed algorithm (Algorithm 2 in the main paper) as optimisation iterations progress. We can see from Figure 1 that the batch size, on average, decreases with increasing optimisation iterations. This property of our algorithm is significant as it minimizes experimental cost by avoiding unnecessary experiments. The optimiser starts with using higher batch size, however it increasingly figures out the function landscape and starts to save on the number of experiments.

4.2 Comparison with other methods

During the review process, we were asked to compare between our proposed method with the filtering of Eq (7) (that is, using $\mathbb{X}_t^+$) or without it (that is, using $\mathbb{X}$). We have performed this comparison using all three benchmark functions used in our paper and present in Figure 2. As seen from the figure, the optimisation performance of our algorithm using $\mathbb{X}_t^+$ is comparable to that of the variant using $\mathbb{X}$. For the Hartmann3 function, the algorithm using $\mathbb{X}$ works better compared to the algorithm using $\mathbb{X}_t^+$. For the Eggholder function, however, the opposite is true. Both algorithms perform similarly for the Branin function. Therefore it is not easy to conclude if one is better than the other empirically. We have chosen to mainly work with the $\mathbb{X}_t^+$ variant as it is possible to obtain a bound on both total cumulative regret and batch cumulative regret, whereas for the variant using $\mathbb{X}$ we can only bound batch cumulative regret.

We were also asked to check how our method compares to the multi-point EI algorithm (Ginsbourger et al., 2010). We have performed this comparison using the implementation available from the URL (https://github.com/cran/DiceOptim/blob/master/R/qEI.R). The results for the Hartmann3 function are shown in Figure 3. We tried to perform this comparison for the other two benchmark functions (eggholder and branin) but could not get reportable results as

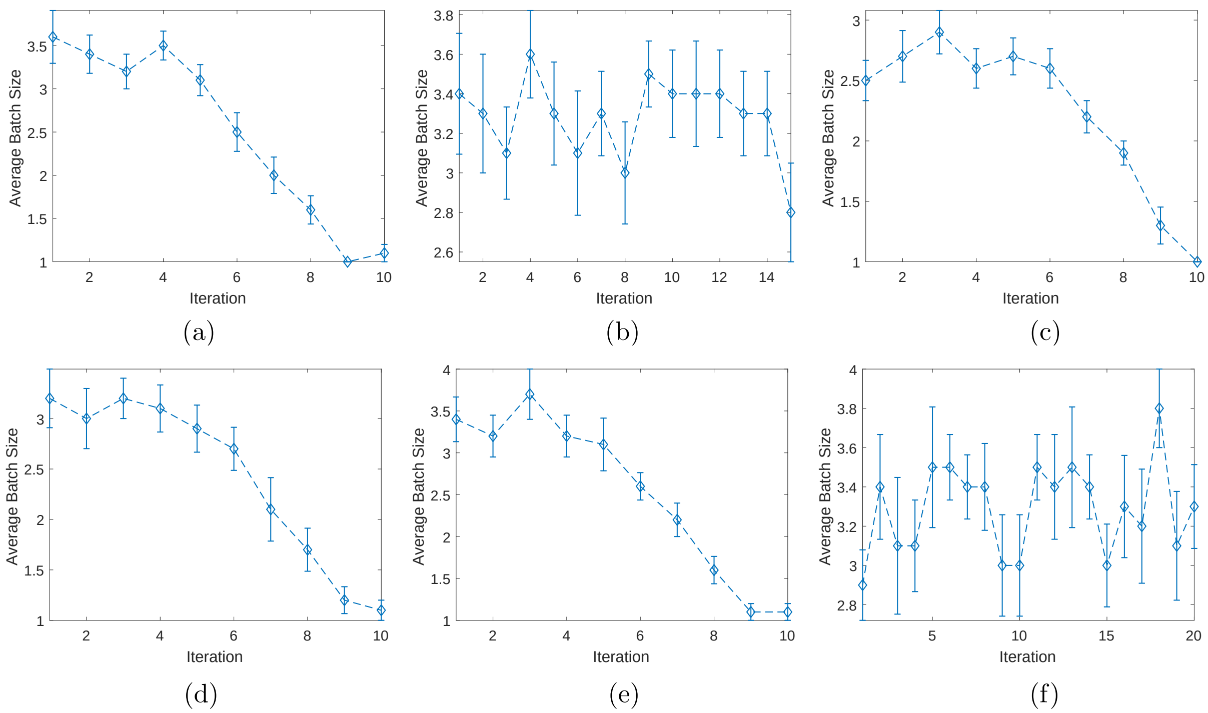

Figure 1: The number of recommendations with respect to optimisation iterations - (a) Branin (b) Hartmann3 (c) Eggholder (d) Support vector regression hyperparameter tuning task (e) Support vector classification hyperparameter tuning task (f) Heat treatment optimisation of Al-Sc alloy.

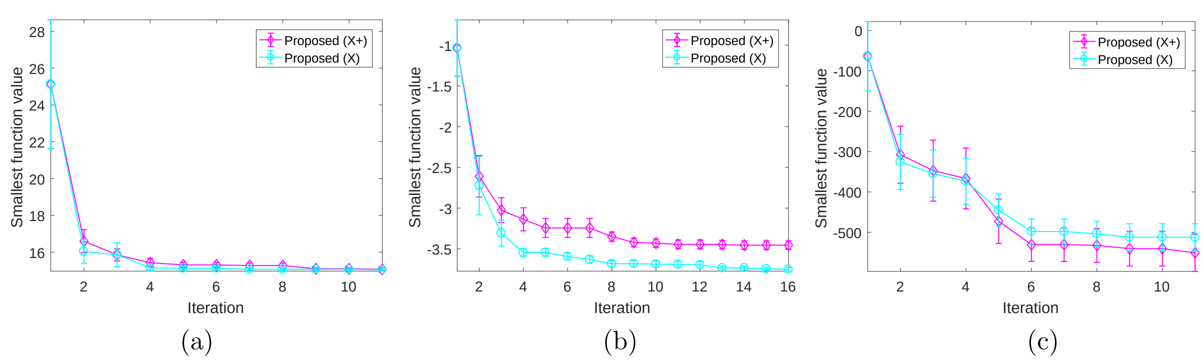

Figure 2: A comparison between two variants of proposed algorithm using $\mathbb{X}_t^+$ or $\mathbb{X}$ - (a) Branin (b) Hartmann3 (c) Eggholder.

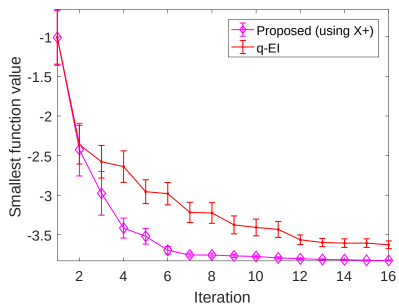

Figure 3: A comparison between our proposed algorithm using $\mathbb{X}_t^+$ and multi-point EI using Hartmann3 function.

the multi-point EI implementation did not improve at all over the initial random points.

References

Emile Contal, David Buffoni, Alexandre Robicquet, and Nicolas Vayatis. Parallel gaussian process optimization with upper confidence bound and pure exploration. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 225–240. Springer, 2013.

David Ginsbourger, Rodolphe Le Riche, and Laurent Carraro. Kriging is well-suited to parallelize optimization. In *Computational Intelligence in Expensive Optimization Problems*, pages 131–162. Springer, 2010.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, May 2012.