Supplementary Materials for the Submission: "Multi-objective Bayesian Optimization using Pareto-frontier Entropy"

A. Proof of Theorem 3.1

The normalization constant Z is written as

$$Z \coloneqq \int_{\mathcal{F}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} = \sum_{m=1}^{M} \int_{\mathcal{C}_m} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} = \sum_{m=1}^{M} \prod_{l=1}^{L} \int_{\ell_m^l}^{u_m^l} p(f_{\boldsymbol{x}}^l \mid \mathcal{D}) \mathrm{d}f_{\boldsymbol{x}}^l = \sum_{m=1}^{M} Z_m. \tag{7}$$

which is a sum of the Gaussian integrals in the cells.

First, (4) is immediately derived from the independence of L GPRs. Let $\mathcal{F} \coloneqq \{\boldsymbol{f}_{\boldsymbol{x}} \mid \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*\}$. Using

$$p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \in \mathcal{F}) = \begin{cases} \frac{1}{Z} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}), & \text{if } \boldsymbol{f}_{\boldsymbol{x}} \in \mathcal{F}, \\ 0, & \text{otherwise ,} \end{cases}$$

we see

$$\begin{aligned} H[p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \in \mathcal{F})] &= -\int_{\mathcal{F}} \frac{p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D})}{Z} \log \frac{p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D})}{Z} \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} \\ &= -Z^{-1} \int_{\mathcal{F}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \log p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} + Z^{-1} \log Z \int_{\mathcal{F}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} \\ &= -Z^{-1} \int_{\mathcal{F}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \log p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} + \log Z \\ &= -Z^{-1} \sum_{m=1}^{M} \int_{\mathcal{C}_m} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \log p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} + \log Z \end{aligned} \tag{8}$$

Based on the independence of $\boldsymbol{f}_{\boldsymbol{x}}$, the integral of the first term can be transformed into

$$\begin{aligned} &\sum_{m=1}^{M} \int_{\mathcal{C}_m} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \log p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}} \\ &= \sum_{m=1}^{M} \int_{\ell_m^1}^{u_m^1} \int_{\ell_m^2}^{u_m^2} \cdots \int_{\ell_m^L}^{u_m^L} \prod_{l'=1}^{L} p(f_{\boldsymbol{x}}^{l'} \mid \mathcal{D}) \left(\sum_{l=1}^{L} \log p(f_{\boldsymbol{x}}^l \mid \mathcal{D}) \right) \mathrm{d}f_{\boldsymbol{x}}^L \cdots \mathrm{d}f_{\boldsymbol{x}}^1 \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L} \int_{\ell_m^1}^{u_m^1} \int_{\ell_m^2}^{u_m^2} \cdots \int_{\ell_m^L}^{u_m^L} \prod_{l'=1}^{L} p(f_{\boldsymbol{x}}^{l'} \mid \mathcal{D}) \log p(f_{\boldsymbol{x}}^l \mid \mathcal{D}) \mathrm{d}f_{\boldsymbol{x}}^L \cdots \mathrm{d}f_{\boldsymbol{x}}^1 \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L} \left[\left(\int_{\ell_m^l}^{u_m^l} p(f_{\boldsymbol{x}}^l \mid \mathcal{D}) \log p(f_{\boldsymbol{x}}^l \mid \mathcal{D}) \mathrm{d}f_{\boldsymbol{x}}^l \right) \left(\prod_{l' \neq l} \int_{\ell_m^{l'}}^{u_m^{l'}} p(f_{l'}(\boldsymbol{x}) \mid \mathcal{D}) \mathrm{d}f_{l'}(\boldsymbol{x}) \right) \right] \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L} \left[Z_{ml} \int_{\ell_m^l}^{u_m^l} \left(\frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \log \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} + \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \log Z_{ml} \right) \mathrm{d}f_{\boldsymbol{x}}^l \prod_{l' \neq l} Z_{ml'} \right] \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L} \left[\left(Z_{ml} \underbrace{\int_{\ell_m^l}^{u_m^l} \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \log \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \mathrm{d}f_{\boldsymbol{x}}^l}_{(\star)} + Z_{ml} \log Z_{ml} \right) \prod_{l' \neq l} Z_{ml'} \right] \end{aligned} \tag{9}$$

The term indicated by $\star$ is the negative entropy of the truncated normal distribution. For the entropy of the truncated normal distribution, analytical formula is available (e.g, Michalowicz et al., 2013), by which we can obtain

$$\int_{\ell_m^l}^{u_m^l} \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \log \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z_{ml}} \mathrm{d} f_{\boldsymbol{x}}^l = -\log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x}) Z_{ml}) - \frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2Z_{ml}}.$$

Then, the above equation (9) is further transformed into

$$\begin{aligned}
&\sum_{m=1}^{M}\sum_{l=1}^{L}\left[\left(Z_{ml}\Big(-\log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})Z_{ml}) - \frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2Z_{ml}}\Big) + Z_{ml}\log Z_{ml}\right)\prod_{l'\neq l} Z_{ml'}\right] \\
&= \sum_{m=1}^{M}\sum_{l=1}^{L}\left[\left(-Z_{ml}\log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})) - \frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2}\right)\prod_{l'\neq l} Z_{ml'}\right].
\end{aligned}$$

Substituting this into (8), we obtain

$$\begin{aligned}
&H[p(\boldsymbol{f_x} \mid \mathcal{D}, \boldsymbol{f_x} \in \mathcal{F})] \\
&= -Z^{-1}\sum_{m=1}^{M}\sum_{l=1}^{L}\left[\left(-Z_{ml}\log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})) - \frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2}\right)\prod_{l'\neq l} Z_{ml'}\right] + \log Z \\
&= Z^{-1}\sum_{m=1}^{M}\prod_{l'=1}^{L} Z_{ml'}\sum_{l=1}^{L}\log\left(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})\right) + Z^{-1}\sum_{m=1}^{M}\sum_{l=1}^{L}\frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2}\prod_{l'\neq l} Z_{ml'} + \log Z \\
&= \log\left((\sqrt{2\pi e})^L Z\prod_{l=1}^{L}\sigma_l(\boldsymbol{x})\right) + Z^{-1}\sum_{m=1}^{M}\sum_{l=1}^{L}\prod_{l'\neq l} Z_{ml'}\frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2} \\
&= \log\left((\sqrt{2\pi e})^L Z\prod_{l=1}^{L}\sigma_l(\boldsymbol{x})\right) + \sum_{m=1}^{M}\frac{Z_m}{Z}\sum_{l=1}^{L}\frac{\underline{\alpha}_{m,l}\phi(\underline{\alpha}_{m,l}) - \bar{\alpha}_{m,l}\phi(\bar{\alpha}_{m,l})}{2Z_{ml}}.
\end{aligned}$$

B. Proof of Theorem 4.1

The marginalization can be represented as

$$\begin{aligned}
&p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f_x} \preceq \mathcal{F}^*) \\
&= \sum_{m \in \mathcal{M}(l, s_l^{(f_{\boldsymbol{x}}^l)})} \int_{\mathcal{C}_m^{\setminus l}} \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z} \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l} \\
&= \frac{p(f_{\boldsymbol{x}}^l \mid \mathcal{D})}{Z} \sum_{m \in \mathcal{M}(l, s_l^{(f_{\boldsymbol{x}}^l)})} \int_{\mathcal{C}_m^{\setminus l}} p(\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l} \mid f_{\boldsymbol{x}}^l, \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l},
\end{aligned} \tag{10}$$

where $\mathcal{C}_m^{\setminus l}$ is the $(L-1)$-dimensional cell created by eliminating the l-th dimension of $\mathcal{C}_m$, and $\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l}$ is a subvector of $\boldsymbol{f_x}$ without the l-th dimension.

The marginal distribution of $f_{\boldsymbol{x}}^l$ can be partitioned into an interval $f_{\boldsymbol{x}}^l \in (\tilde{f}_l^m, \tilde{f}_l^{m+1}]$ as shown in (10), which can be further transformed into

$$
\begin{aligned}
&p(f^l_{\boldsymbol{x}} \in (\tilde{f}^s_l, \tilde{f}^{s+1}_l] \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \\
&= \sum_{m' \in \mathcal{M}(l, s_l^{(f^l_{\boldsymbol{x}})})} \int_{\mathcal{C}^{\setminus l}_{m'}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \mathrm{d}\boldsymbol{f}_{\setminus l}(\boldsymbol{x}) \\
&= \frac{1}{Z} \sum_{m' \in \mathcal{M}(l, s_l^{(f^l_{\boldsymbol{x}})})} \int_{\mathcal{C}^{\setminus l}_{m'}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\setminus l}(\boldsymbol{x}) \\
&= \frac{1}{Z} \sum_{m' \in \mathcal{M}(l, s_l^{(f^l_{\boldsymbol{x}})})} \int_{\mathcal{C}^{\setminus l}_{m'}} \prod_{l' \neq l} p(f_{l'}(\boldsymbol{x}) \mid \mathcal{D}) p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\setminus l}(\boldsymbol{x}) \\
&= \frac{1}{Z} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \sum_{m' \in \mathcal{M}(l, s_l^{(f^l_{\boldsymbol{x}})})} \prod_{l' \neq l} (\Phi(\bar{\alpha}_{m',l'}) - \Phi(\underline{\alpha}_{m',l'})) \\
&= \frac{\sum_{m' \in \mathcal{M}(l, s_l^{(f^l_{\boldsymbol{x}})})} \prod_{l' \neq l} Z_{m'l'}}{Z} p(f^l_{\boldsymbol{x}} \mid \mathcal{D})
\end{aligned}
$$

Let $\tilde{f}^0_l = -\infty$, for convenience. Then, the entropy is

$$
\begin{aligned}
&H[p(f^l_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)] \\
&= -\int_{-\infty}^{\infty} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \log p(f^l_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= \sum_{s=0}^{S_l - 1} \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \log p(f^l_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= -\sum_{s=0}^{S_l - 1} \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \log \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= -\sum_{s=0}^{S_l - 1} \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \left(\log \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} + \log p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \right) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= -\sum_{s=0}^{S_l - 1} \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \Bigg((\Phi(\tilde{\alpha}_{s+1,l}) - \Phi(\tilde{\alpha}_{s,l})) \log \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \\
&\quad + \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \log p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}f^l_{\boldsymbol{x}} \Bigg) \qquad (11)
\end{aligned}
$$

By transforming the last term in the parenthesis into the entropy of the truncated normal distribution, we see

$$
\begin{aligned}
&\int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \log p(f^l_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= \tilde{Z}_{sl} \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} \frac{p(f^l_{\boldsymbol{x}} \mid \mathcal{D})}{\tilde{Z}_{sl}} \left(\log \frac{p(f^l_{\boldsymbol{x}} \mid \mathcal{D})}{\tilde{Z}_{sl}} + \log \tilde{Z}_{sl} \right) \mathrm{d}f^l_{\boldsymbol{x}} \\
&= -\tilde{Z}_{sl} \left\{ \log(\sqrt{2\pi e} \sigma_l(\boldsymbol{x}) \tilde{Z}_{sl}) + \frac{\tilde{\alpha}_{s,l} \phi(\tilde{\alpha}_{s,l}) - \tilde{\alpha}_{s+1,l} \phi(\tilde{\alpha}_{s+1,l})}{2\tilde{Z}_{sl}} \right\} \\
&\quad + \tilde{Z}_{sl} \int_{\tilde{f}^s_l}^{\tilde{f}^{s+1}_l} \frac{p(f^l_{\boldsymbol{x}} \mid \mathcal{D})}{\tilde{Z}_{sl}} \log \tilde{Z}_{sl} \mathrm{d}f^l_{\boldsymbol{x}} \\
&= -\tilde{Z}_{sl} \left\{ \log(\sqrt{2\pi e} \sigma_l(\boldsymbol{x}) \tilde{Z}_{sl}) + \frac{\tilde{\alpha}_{s,l} \phi(\tilde{\alpha}_{s,l}) - \tilde{\alpha}_{s+1,l} \phi(\tilde{\alpha}_{s+1,l})}{2\tilde{Z}_{sl}} \right\} + \tilde{Z}_{sl} \log \tilde{Z}_{sl} \qquad (12)
\end{aligned}
$$

By substituting this into (11), we obtain

$$
\begin{aligned}
&H[p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)] \\
&= -\sum_{s=0}^{S_l-1} \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \left(\tilde{Z}_{sl} \log \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \right. \\
&\quad \left. - \tilde{Z}_{sl} \left\{ \log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})\tilde{Z}_{sl}) + \frac{\tilde{\alpha}_{s,l}\phi(\tilde{\alpha}_{s,l}) - \tilde{\alpha}_{s+1,l}\phi(\tilde{\alpha}_{s+1,l})}{2\tilde{Z}_{sl}} \right\} + \tilde{Z}_{sl} \log \tilde{Z}_{sl} \right) \\
&= -\sum_{s=0}^{S_l-1} \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \tilde{Z}_{sl} \left(\log \frac{\sum_{m' \in \mathcal{M}(l,s)} \prod_{l' \neq l} Z_{m'l'}}{Z} \right. \\
&\quad \left. - \log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})\tilde{Z}_{sl}) - \frac{\tilde{\alpha}_{s,l}\phi(\tilde{\alpha}_{s,l}) - \tilde{\alpha}_{s+1,l}\phi(\tilde{\alpha}_{s+1,l})}{2\tilde{Z}_{sl}} + \log \tilde{Z}_{sl} \right)
\end{aligned}
$$

From the definition, if $m' \in \mathcal{M}(l,s)$, then $\tilde{Z}_{sl} = Z_{m'l}$, from which we obtain $\sum_{m' \in \mathcal{M}(l,s)} \left(\prod_{l' \neq l} Z_{m'l'} \right) \tilde{Z}_{sl} = \sum_{m' \in \mathcal{M}(l,s)} \left(\prod_{l'=1}^{L} Z_{m'l'} \right) = \sum_{m' \in \mathcal{M}(l,s)} Z_{m'}$. This derives

$$
\begin{aligned}
&H[p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)] \\
&= -\sum_{s=0}^{S_l-1} \frac{\sum_{m' \in \mathcal{M}(l,s)} Z_{m'}}{Z} \left(\log \frac{\sum_{m' \in \mathcal{M}(l,s)} Z_{m'}}{Z} \right. \\
&\quad \left. - \log(\sqrt{2\pi e}\sigma_l(\boldsymbol{x})\tilde{Z}_{sl}) - \frac{\tilde{\alpha}_{s,l}\phi(\tilde{\alpha}_{s,l}) - \tilde{\alpha}_{s+1,l}\phi(\tilde{\alpha}_{s+1,l})}{2\tilde{Z}_{sl}} \right)
\end{aligned}
$$

C. Extension to Correlated Objectives

Objective functions in MOO are often correlated each other. Then, by incorporating the correlation into GPR, the search can be accelerated. Several studies have considered constructing multiple correlated GPR models including *multi-task GPR* model (Bonilla et al., 2008) and *semiparametric latent factor* (SLF) model (Seeger et al., 2004). In the standard approaches including multi-task GPR and SLF, the multi-dimensional predictive distribution for $\boldsymbol{x}$ is reduced to a multi-variate Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{x}), \boldsymbol{\Sigma}(\boldsymbol{x}))$, where $\boldsymbol{\mu}(\boldsymbol{x}) \in \mathbb{R}^L$ and $\boldsymbol{\Sigma}(\boldsymbol{x}) \in \mathbb{R}^{L \times L}$ are the predictive mean and covariance matrix. For considering an extension of PFES to correlated objectives, we assume that the surrogate model is represented as a GPR model jointly for multiple responses.

For the coupled setting, we need to evaluate analytically intractable integrations in (7) and (8). The normalization constant Z (7) is defined by the sum of the integral of Gaussian distribution on the hyper-rectangle region ($\mathcal{C}_m$). The numerical computation of this form of integrations have been extensively studied (Genz & Bretz, 2009) mainly in the context of the Gaussian probability calculation. The integration in the entropy (8) can also be evaluated through the Gaussian probability (Appendix D shows computational detail). Although this approach requires $O(L)$ times $L-1$-dimensional and $O(L^2)$ times $L-2$-dimensional CDF calculations, in many practical problems, the number of objectives L is quite small.

For the decoupled setting, if $L = 2$, we can derive a simple form of the entropy calculation because the conditional distribution $p(\boldsymbol{f}_{\boldsymbol{x}}^{\backslash l} \mid f_{\boldsymbol{x}}^l, \mathcal{D})$ in (10) becomes a one-dimensional Gaussian distribution. Let $\sigma_{12}^2(\boldsymbol{x})$ be the predictive covariance of two-dimensional $\boldsymbol{f}_{\boldsymbol{x}} = (f_{\boldsymbol{x}}^1, f_{\boldsymbol{x}}^2)^\top$. Then, we obtain the following theorem:

Theorem C.1. *Let* $W_{m'2}(f_{\boldsymbol{x}}^1) \coloneqq \Phi\left(\frac{f_2^{m'+1} - u(\boldsymbol{x} \mid f_{\boldsymbol{x}}^1)}{s(\boldsymbol{x})}\right) - \Phi\left(\frac{f_2^{m'} - u(\boldsymbol{x} \mid f_{\boldsymbol{x}}^1)}{s(\boldsymbol{x})}\right)$*, where* $u(\boldsymbol{x} \mid f) \coloneqq \frac{\sigma_{12}^2(\boldsymbol{x})(f - \mu_1(\boldsymbol{x}))}{\sigma_1^2(\boldsymbol{x})} + \mu_2(\boldsymbol{x})$ *and* $s^2(\boldsymbol{x}) \coloneqq \sigma_2^2(\boldsymbol{x}) - \frac{(\sigma_{12}^2(\boldsymbol{x}))^2}{\sigma_1^2(\boldsymbol{x})}$*. For the two dimensional correlated GPRs* $\boldsymbol{f}(\boldsymbol{x})$*, the entropy of* $p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)$ *is given by*

$$
H[p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)] = -\sum_{s=0}^{S_1-1} \int_{\tilde{f}_1^s}^{\tilde{f}_1^{s+1}} \left(\frac{\phi(\alpha_{\boldsymbol{x}}^1)}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1) \right) \log \left(\frac{\phi(\alpha_{\boldsymbol{x}}^1)}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1) \right) \mathrm{d}f_{\boldsymbol{x}}^1
$$

where $\alpha_{\boldsymbol{x}}^1 \coloneqq (f_{\boldsymbol{x}}^1 - \mu_1(\boldsymbol{x}))/\sigma_1(\boldsymbol{x})$.

Proof. Let $u(\boldsymbol{x} \mid f) \coloneqq \frac{\sigma_{12}^2(\boldsymbol{x})(f-\mu_1(\boldsymbol{x}))}{\sigma_1^2(\boldsymbol{x})} + \mu_2(\boldsymbol{x})$ and $s^2(\boldsymbol{x}) \coloneqq \sigma_2^2(\boldsymbol{x}) - \frac{(\sigma_{12}^2(\boldsymbol{x}))^2}{\sigma_1^2(\boldsymbol{x})}$. From (10), the marginal can be written as follows:

$$
\begin{aligned}
p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) &= \frac{p(f_{\boldsymbol{x}}^1 \mid \mathcal{D})}{Z} \sum_{m' \in \mathcal{M}(1, s_1^{(f_{\boldsymbol{x}}^1)})} \int_{\ell_{m'}^2}^{u_{m'}^2} p(f_{\boldsymbol{x}}^2 \mid f_{\boldsymbol{x}}^1, \mathcal{D}) \mathrm{d} f_{\boldsymbol{x}}^2 \\
&= \frac{p(f_{\boldsymbol{x}}^1 \mid \mathcal{D})}{Z} \sum_{m' \in \mathcal{M}(1, s_1^{(f_{\boldsymbol{x}}^1)})} W_{m'2}(f_{\boldsymbol{x}}^1)
\end{aligned}
$$

where

$$
W_{m'2}(f_{\boldsymbol{x}}^1) \coloneqq \Phi\left(\frac{u_{m'}^2 - u(\boldsymbol{x} \mid f_{\boldsymbol{x}}^1)}{s(\boldsymbol{x})}\right) - \Phi\left(\frac{\ell_{m'}^2 - u(\boldsymbol{x} \mid f_{\boldsymbol{x}}^1)}{s(\boldsymbol{x})}\right)
$$

Then, the entropy is

$$
\begin{aligned}
&H[p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)] \\
&= -\sum_{s=0}^{S_1-1} \int_{\tilde{f}_1^s}^{\tilde{f}_1^{s+1}} p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \log p(f_{\boldsymbol{x}}^1 \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \mathrm{d} f_{\boldsymbol{x}}^1 \\
&= -\sum_{s=0}^{S_1-1} \int_{\tilde{f}_1^s}^{\tilde{f}_1^{s+1}} \left(\frac{p(f_{\boldsymbol{x}}^1 \mid \mathcal{D})}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1)\right) \log \left(\frac{p(f_{\boldsymbol{x}}^1 \mid \mathcal{D})}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1)\right) \mathrm{d} f_{\boldsymbol{x}}^1 \\
&= -\sum_{s=0}^{S_1-1} \int_{\tilde{f}_1^s}^{\tilde{f}_1^{s+1}} \left(\frac{\phi(\alpha_{\boldsymbol{x}}^1)}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1)\right) \log \left(\frac{\phi(\alpha_{\boldsymbol{x}}^1)}{Z} \sum_{m' \in \mathcal{M}(1,s)} W_{m'2}(f_{\boldsymbol{x}}^1)\right) \mathrm{d} f_{\boldsymbol{x}}^1
\end{aligned}
$$

where $\alpha_{\boldsymbol{x}}^1 \coloneqq (f_{\boldsymbol{x}}^1 - \mu_1(\boldsymbol{x}))/\sigma_1(\boldsymbol{x})$. □

Although the integral inside the sum is analytically intractable, we can numerically calculate it easily because the integral is over the one-dimensional interval.

In the case of $L > 2$, the integral $\int_{\mathcal{C}_{m'}^{\setminus l}} p(\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l} \mid f_{\boldsymbol{x}}^l, \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l}$ in (10) is also the multi-dimensional Gaussian integration (Genz & Bretz, 2009). The marginal density $p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*)$ defined by (10) can also be evaluated through the integration of the Gaussian density because $p(\boldsymbol{f}_{\boldsymbol{x}}^{\setminus l} \mid f_{\boldsymbol{x}}^l, \mathcal{D})$ can be analytically derived for a given $f_{\boldsymbol{x}}^l$. Here again, for the integral in (10), we can use numerical technique for the Gaussian probability (Genz & Bretz, 2009). Then, we can simply approximate the integral of the entropy $\int_{f_{\boldsymbol{x}}^l} p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \log p(f_{\boldsymbol{x}}^l \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \preceq \mathcal{F}^*) \mathrm{d} f_{\boldsymbol{x}}^l$ by a sum of finite grid points. This is also one-dimensional integral, and thus accurate approximation can be expected.

D. Entropy Evaluation for Correlated Objectives

Here, we redefine

$$
Z \coloneqq \int_{\mathcal{F}} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}}, \text{ and } Z_m \coloneqq \int_{\mathcal{C}_m} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}},
$$

which indicate $Z = \sum_{m=1}^{M} Z_m$. The entropy of the conditional distribution $p(\boldsymbol{f_x} \mid \mathcal{D}, \boldsymbol{f_x} \preceq \mathcal{F}^*)$ can be transformed as follows:

$$\begin{aligned} H[p(\boldsymbol{f_x} \mid \mathcal{D}, \boldsymbol{f_x} \preceq \mathcal{F}^*)] &= -\int_{\mathcal{F}} \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z} \log \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z} \mathrm{d}\boldsymbol{f_x} \\ &= -\frac{1}{Z} \sum_{m=1}^{M} \int_{\mathcal{C}_m} p(\boldsymbol{f_x} \mid \mathcal{D}) \log p(\boldsymbol{f_x} \mid \mathcal{D}) \mathrm{d}\boldsymbol{f_x} + \log Z \\ &= \frac{1}{Z} \sum_{m=1}^{M} Z_m \left(\underbrace{-\int_{\mathcal{C}_m} \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z_m} \log \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z_m} \mathrm{d}\boldsymbol{f_x} + \log Z_m}_{(\star)} \right) - \log Z \end{aligned} \tag{13}$$

The term indicated by $\star$ is the entropy of the multi-variate truncated normal distribution. This term can also be written as $-\mathbb{E}_{\mathrm{TN}} \left[\log \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z_m} \right]$, where $\mathbb{E}_{\mathrm{TN}}$ is an expectation by the truncated normal distribution

$$p(\boldsymbol{f_x} \mid \mathcal{D}, \boldsymbol{f_x} \in \mathcal{C}_m) = \begin{cases} \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})/Z_m, & \text{if } \boldsymbol{f_x} \in \mathcal{C}_m, \\ 0, & \text{otherwise,} \end{cases}$$

with the predictive mean $\boldsymbol{\mu} \in \mathbb{R}^L$ and the predictive covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{L \times L}$ of the current GPR. We derive that this entropy can be represented through the moment of the truncated normal distribution:

$$\begin{aligned} -\mathbb{E}_{\mathrm{TN}} \left[\log \frac{p(\boldsymbol{f_x} \mid \mathcal{D})}{Z_m} \right] &= -\mathbb{E}_{\mathrm{TN}} \left[\log p(\boldsymbol{f_x} \mid \mathcal{D}) \right] + \log Z_m \\ &= -\mathbb{E}_{\mathrm{TN}} \left[-\frac{1}{2} \log |2\pi \boldsymbol{\Sigma}| - \frac{1}{2} (\boldsymbol{f_x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{f_x} - \boldsymbol{\mu}) \right] + \log Z_m \\ &= \frac{1}{2} \log |2\pi \boldsymbol{\Sigma}| + \frac{1}{2} \mathbb{E}_{\mathrm{TN}} \left[(\boldsymbol{f_x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{f_x} - \boldsymbol{\mu}) \right] + \log Z_m \end{aligned} \tag{14}$$

Let $\boldsymbol{\mu}_{\mathrm{TN}} \coloneqq \mathbb{E}_{\mathrm{TN}}[\boldsymbol{f_x}]$, $\boldsymbol{d} \coloneqq \boldsymbol{\mu}_{\mathrm{TN}} - \boldsymbol{\mu}$, and $\boldsymbol{\Sigma}_{\mathrm{TN}} \coloneqq \mathbb{E}_{\mathrm{TN}}[(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}})(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}})^\top]$. Then, the second term of the above equation (14) is written as

$$\begin{aligned} \mathbb{E}_{\mathrm{TN}} \left[(\boldsymbol{f_x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{f_x} - \boldsymbol{\mu}) \right] &= \mathrm{Trace} \left(\boldsymbol{\Sigma}^{-1} \mathbb{E}_{\mathrm{TN}} \left[(\boldsymbol{f_x} - \boldsymbol{\mu})(\boldsymbol{f_x} - \boldsymbol{\mu})^\top \right] \right) \\ &= \mathrm{Trace} \left(\boldsymbol{\Sigma}^{-1} \mathbb{E}_{\mathrm{TN}} \left[(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}} + \boldsymbol{d})(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}} + \boldsymbol{d})^\top \right] \right) \\ &= \mathrm{Trace} \left(\boldsymbol{\Sigma}^{-1} (\mathbb{E}_{\mathrm{TN}} \left[(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}})(\boldsymbol{f_x} - \boldsymbol{\mu}_{\mathrm{TN}})^\top \right] + \mathbb{E}_{\mathrm{TN}} \left[\boldsymbol{d}\boldsymbol{d}^\top \right]) \right) \\ &= \mathrm{Trace} \left(\boldsymbol{\Sigma}^{-1} (\boldsymbol{\Sigma}_{\mathrm{TN}} + \boldsymbol{d}\boldsymbol{d}^\top) \right) \end{aligned} \tag{15}$$

If $\boldsymbol{\mu}_{\mathrm{TN}}$ and $\boldsymbol{\Sigma}_{\mathrm{TN}}$ are available, the entropy (13) can be evaluated by combining (14) and (15).

The two expected values $\boldsymbol{\mu}_{\mathrm{TN}}$ and $\boldsymbol{\Sigma}_{\mathrm{TN}}$ can be obtained from the first and the second moment of the multi-variate truncated normal distribution, for which Manjunath & Wilhelm (2009) show efficient computations through the Gaussian integral calculation. Let μ^i and μ^i_{TN} be the i-th element of $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_{\mathrm{TN}}$, respectively, and let $\sigma^{i,j}$ and $\sigma^{i,j}_{\mathrm{TN}}$ be the (i,j)-th element of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_{\mathrm{TN}}$, respectively. By defining the k-th dimensional marginal distribution of the truncated normal as

$$\begin{aligned} F_k(f^k_{\boldsymbol{x}}) &= p(f^k_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f_x} \in \mathcal{C}_m) \\ &= \int_{\ell^1_m}^{u^1_m} \cdots \int_{\ell^{k-1}_m}^{u^{k-1}_m} \int_{\ell^{k+1}_m}^{u^{k+1}_m} \cdots \int_{\ell^L_{m+1}}^{u^L_m} p(\boldsymbol{f_x} \mid \mathcal{D}, \boldsymbol{f_x} \in \mathcal{C}_m) \mathrm{d}\boldsymbol{f}^{\setminus k}_{\boldsymbol{x}}, \end{aligned}$$

μ^i_{TN} can be represented as

$$\begin{aligned} \mu^i_{\mathrm{TN}} &= \mu^i + \sum_{k=1}^{L} \sigma^{i,k} \big(F_k(\ell^k_m) - F_k(u^k_m) \big), \\ &= \mu^i + d^i, \end{aligned} \tag{16}$$

Table 2: Computational time of acquisition function on Ackley/Sphere 2D.

		50 train data	100 train data	200 train data
ParEGO		0.44 ± 0.01	0.67 ± 0.02	1.32 ± 0.07
SMSego		0.20 ± 0.01	0.35 ± 0.01	0.62 ± 0.02
EHI		0.10 ± 0.00	0.13 ± 0.00	0.13 ± 0.00
MESMO		48.35 ± 1.27	48.41 ± 0.85	48.22 ± 0.66
PFES	Total	39.42 ± 0.65	41.20 ± 0.54	43.81 ± 0.86
	RFM	0.05 ± 0.00	0.060 ± 0.00	0.06 ± 0.00
	NSGAII	38.97 ± 0.67	40.67 ± 0.56	43.26 ± 0.88
	QHV	0.040 ± 0.00	0.05 ± 0.00	0.04 ± 0.00
	eval entropy	0.36 ± 0.02	0.43 ± 0.03	0.44 ± 0.02
	# Cell	50.00 ± 0.00	50.00 ± 0.00	50.00 ± 0.00

where $d^i \coloneqq \sum_{k=1}^{L} \sigma^{i,k} \big(F_k(\ell_m^k) - F_k(u_m^k)\big)$. For $\sigma_{\rm TN}^{i,j}$, by defining the (k,q)-th two dimensional marginal distribution of the truncated normal as

$$
\begin{aligned}
F_{k,q}(f_{\boldsymbol{x}}^k, f_{\boldsymbol{x}}^q) &= p(f_{\boldsymbol{x}}^k, f_{\boldsymbol{x}}^q \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \in \mathcal{C}_m) \\
&= \int_{\ell_m^1}^{u_m^1} \cdots \int_{\ell_m^{k-1}}^{u_m^{k-1}} \int_{\ell_m^{k+1}}^{u_m^{k+1}} \cdots \int_{\ell_m^{q-1}}^{u_m^{q-1}} \int_{\ell_m^{q+1}}^{u_m^{q+1}} \cdots \int_{\ell_m^L}^{u_m^L} p(\boldsymbol{f}_{\boldsymbol{x}} \mid \mathcal{D}, \boldsymbol{f}_{\boldsymbol{x}} \in \mathcal{C}_m) \mathrm{d}\boldsymbol{f}_{\boldsymbol{x}}^{\setminus k,q},
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\sigma_{TN}^{i,j} = \sigma^{i,j} - d^i d^j &+ \sum_{k=1}^{L} \sigma^{i,k} \frac{\sigma^{j,k}\big(\ell_m^k F_k(\ell_m^k) - u_k^m F_k(u_m^k)\big)}{\sigma^{k,k}} \\
&+ \sum_{k=1}^{L} \sigma^{i,k} \sum_{q \neq k} \left(\sigma^{j,q} - \frac{\sigma^{k,q}\sigma^{j,k}}{\sigma^{k,k}}\right) \\
&\quad \cdot \left[\big(F_{k,q}(\ell_m^k, \ell_m^q) - F_{k,q}(\ell_m^k, u_m^q)\big) - \big(F_{k,q}(u_m^k, \ell_m^q) - F_{k,q}(u_m^k, u_m^q)\big)\right].
\end{aligned} \tag{17}
$$

Thus, to calculate (16) and (17), $O(L)$ times $L-1$ dimensional Gaussian integration and $O(L^2)$ times $L-2$ dimensional Gaussian integration are necessary.

E. Acquisition Function Computation

We randomly selected $50, 100$, and 200 training instances, and calculated each acquisition function for randomly selected 100 points. We measured CPU time on our python code by the single thread execution. Precise evaluation of computational cost is difficult because of its dependence on implementation detail. Our main purpose here is to show PFES is feasible enough for reasonable size of L. The results are shown in Table 2-5 (OOM indicates out-of-memory).

Overall, the results have the same tendency as Table 1 in our main text. Note that since Ackley/Sphere 2D and ZDT4 are $L = 2$, #cells in PFES is always 50, which is equal to $|\mathrm{PF}|$.

F. Experimental Settings

For GPR, we used GPy (`https://sheffieldml.github.io/GPy/`). The noise term in GPR $\sigma_{\rm noise}$ is fixed at 10^{-4}. The marginal likelihood optimization of the Gaussian kernel parameter $\sigma > 0$ is performed by using gradient descent. This optimization was performed at every iteration in the benchmark experiments. For the material datasets, we first randomly selected 100 samples to optimize σ and $\sigma_{\rm noise}$ through the marginal likelihood optimization. These values were fixed during the BO procedure. Since the material datasets are noisy, we employed this approach for avoiding unstable behaviors of all the compared methods because of the unstable GPR hyper-parameters. We implemented ParEGO and SMSego by ourselves. For ParEGO, the weighting constant ρ of the augmented Tchebycheff function is set 0.05

Table 3: Computational time of acquisition function on ZDT4.

		50 train data	100 train data	200 train data
ParEGO		0.40 ± 0.05	0.55 ± 0.06	1.23 ± 0.02
SMSego		0.24 ± 0.02	0.37 ± 0.03	0.71 ± 0.03
EHI		0.11 ± 0.00	0.11 ± 0.00	0.13 ± 0.00
MESMO		51.33 ± 0.47	48.40 ± 0.52	44.17 ± 0.63
PFES	Total	51.59 ± 0.44	48.71 ± 0.35	51.89 ± 0.16
	RFM	0.06 ± 0.00	0.06 ± 0.00	0.07 ± 0.00
	NSGAII	50.96 ± 0.44	48.13 ± 0.37	51.26 ± 0.16
	QHV	0.05 ± 0.00	0.05 ± 0.00	0.05 ± 0.00
	eval entropy	0.51 ± 0.02	0.47 ± 0.03	0.51 ± 0.01
	# Cell	50.00 ± 0.00	50.00 ± 0.00	50.00 ± 0.00

Table 4: Computational time of acquisition function on DTLZ3

		50 train data	100 train data	200 train data
ParEGO		0.51 ± 0.01	0.68 ± 0.03	1.35 ± 0.04
SMSego		1.58 ± 0.61	2.36 ± 0.90	2.75 ± 0.61
EHI		6.26 ± 5.76	21.19 ± 30.11	20.21 ± 11.03
MESMO		62.88 ± 1.39	49.38 ± 0.72	62.25 ± 0.93
PFES	Total	65.36 ± 2.02	57.73 ± 1.26	64.97 ± 1.39
	RFM	0.11 ± 0.00	0.11 ± 0.00	0.13 ± 0.00
	NSGAII	63.024 ± 2.07	55.35 ± 1.25	62.23 ± 1.40
	QHV	1.12 ± 0.12	1.15 ± 0.16	1.32 ± 0.09
	eval entropy	1.10 ± 0.09	1.12 ± 0.07	1.29 ± 0.08
	# Cell	570.99 ± 87.26	614.56 ± 109.27	662.69 ± 107.33

as indicated by the original paper (Knowles, 2006). For SMSego, the coefficient of lower confidence bound is set as $\beta_t = \Phi^{-1}(0.5 + 1/2^L)$ which is also indicated by the original paper (Ponweiser et al., 2008). For MESMO, about Pareto sampling, we used the same settings as PFES. The number of sampling is 10, and the number of basis of RFM is 500.

Table 5: Computational time of acquisition function on DTLZ4

		50 train data	100 train data	200 train data
ParEGO		0.53 ± 0.01	0.64 ± 0.02	1.44 ± 0.04
SMSego		6.32 ± 1.50	11.83 ± 1.92	19.86 ± 2.89
EHI		317.55 ± 63.18	796.10 ± 199.66	OOM
MESMO		59.90 ± 0.43	62.89 ± 0.61	65.30 ± 0.93
PFES	Total	62.36 ± 0.72	60.97 ± 0.36	67.85 ± 0.86
	RFM	0.11 ± 0.00	0.11 ± 0.00	0.14 ± 0.00
	NSGAII	59.85 ± 0.79	58.39 ± 0.35	64.77 ± 0.97
	QHV	1.28 ± 0.08	1.23 ± 0.08	1.59 ± 0.24
	eval entropy	1.13 ± 0.07	1.25 ± 0.04	1.35 ± 0.12
	# Cell	647.86 ± 124.84	645.88 ± 116.01	710.44 ± 141.77