A Proofs

Lemma 4. *Let $g(\mathcal{S}) = \sum_{x \in Q(\mathcal{S})} f(x)$, where $Q(\mathcal{S})$ is defined in Eq. (3.3). If $\forall x,\ f(x) \geq 0$, then g is monotone supermodular.*

Proof. Let $\ell \in [L]$, and $j \in \mathcal{C}^{(\ell)}$ be any constraint at site ℓ. For $\mathcal{S} \subseteq \mathcal{C} \backslash \{j\}$, define $\Delta_g(j \mid \mathcal{S}) = \sum_{x \in Q(\mathcal{S} \cup \{j\})} f(x) - \sum_{x \in Q(\mathcal{S})} f(x)$ to be the gain of adding j to the set $\mathcal{S}$.

By definition of $Q(\mathcal{S})$, we have $Q(\mathcal{S}) = \prod_{k=1}^{L} \mathcal{S}^{(k)}$, and

$$\begin{aligned} Q(\mathcal{S} \cup \{j\}) &= \left(\mathcal{S}^{(\ell)} \cup \{j\}\right) \times \prod_{k \neq \ell} \mathcal{S}^{(k)} \\ &= \left(\{j\} \times \prod_{k \neq \ell} \mathcal{S}^{(k)}\right) \bigcup \left(\mathcal{S}^{(\ell)} \times \prod_{k \neq \ell} \mathcal{S}^{(k)}\right) \\ &= \left(\{j\} \times \prod_{k \neq \ell} \mathcal{S}^{(k)}\right) \bigcup \left(\prod_{k=1}^{L} \mathcal{S}^{(k)}\right) \end{aligned} \tag{A.1}$$

Then,

$$\Delta_g(j \mid \mathcal{S}) = \sum_{x \in Q(\mathcal{S} \cup \{j\})} f(x) - \sum_{x \in Q(\mathcal{S})} f(x) \overset{Eq.\ (A.1)}{=} \sum_{x \in \{j\} \times \prod_{k \neq \ell} \mathcal{S}^{(k)}} f(x)$$

Now let us consider $\mathcal{S}'$ such that $\mathcal{S} \subseteq \mathcal{S}' \subseteq \mathcal{C} \setminus \{j\}$. Clearly $\forall k \in [L],\ \mathcal{S}^{(k)} \subseteq \mathcal{S}'^{(k)}$. Therefore, $\Delta_g(j \mid \mathcal{S}') - \Delta_g(j \mid \mathcal{S}) = \sum_{x \in \{j\} \times \prod_{k \neq \ell} (\mathcal{S}'^{(k)} \setminus \mathcal{S}^{(k)})} f(x) \geq 0$ and hence g is supermodular. □

A.1 Proof of Lemma 2

We now show that Algorithm 3 leads to a polynomial algorithm for constructing a lower bound on Eq. (4.2), and hence on constructing a DS-decomposition of the surrogate objective function $\hat{F}$ (Eq. (3.2)).

Proof of Lemma 2. Let $g(\mathcal{S}) = \sum_{x \in Q(\mathcal{S})} f(x)$. By definition we have

$$\hat{F}(\mathcal{S}) = g(\mathcal{S}) \left(1 - \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n\right) = \underbrace{g(\mathcal{S})}_{\hat{F}_1(\mathcal{S})} - \underbrace{g(\mathcal{S}) \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n}_{\hat{F}_2(\mathcal{S})} = \hat{F}_1(\mathcal{S}) - \hat{F}_2(\mathcal{S})$$

We know from Lemma 4 that $\hat{F}_1$ is supermodular. Let $j \in \mathcal{C}$ and $\mathcal{S} \subseteq \mathcal{C} \setminus \{j\}$. The gain of j on $\hat{F}_1$, denote by $\Delta_1(j \mid \mathcal{S})$, is monotone decreasing.

Let $\Delta_2(j \mid \mathcal{S}) = \hat{F}_2(\mathcal{S} \cup \{j\}) - \hat{F}_2(\mathcal{S})$. Our goal is to find a lower bound on

$$\begin{aligned} \beta &= \min_{\mathcal{S} \subseteq \mathcal{S}' \subseteq \mathcal{C} \backslash j} \left(\Delta_{\hat{F}}(j \mid \mathcal{S}) - \Delta_{\hat{F}}(j \mid \mathcal{S}')\right) \\ &= \min_{\mathcal{S} \subseteq \mathcal{S}' \subseteq \mathcal{C} \backslash j} \left(\underbrace{\Delta_1(j \mid \mathcal{S}) - \Delta_1(j \mid \mathcal{S}')}_{\geq 0} + \Delta_2(j \mid \mathcal{S}) - \Delta_2(j \mid \mathcal{S}')\right) \end{aligned} \tag{A.2}$$

Therefore, it suffices to find a lower bound $\Delta_2(j \mid \mathcal{S}) - \Delta_2(j \mid \mathcal{S}')$. The gain of j on $\hat{F}_2$ is

$$\begin{aligned}\Delta_2(j \mid \mathcal{S}) &= \hat{F}_2(\mathcal{S} \cup \{j\}) - \hat{F}_2(\mathcal{S}) \\ &= \sum_{x \in Q(\mathcal{S}\cup\{j\})} f(x)\left(1 - \frac{1}{|Q(\mathcal{S}\cup\{j\})|}\right)^n - \sum_{x\in Q(\mathcal{S})} f(x)\left(1-\frac{1}{|Q(\mathcal{S})|}\right)^n \\ &= \sum_{x \in Q(\mathcal{S}\cup\{j\})\setminus Q(\mathcal{S})} f(x)\left(1 - \frac{1}{|Q(\mathcal{S}\cup\{j\})|}\right)^n + \\ &\qquad \sum_{x\in Q(\mathcal{S})} f(x)\left(\left(1 - \frac{1}{|Q(\mathcal{S}\cup\{j\})|}\right)^n - \left(1-\frac{1}{|Q(\mathcal{S})|}\right)^n\right)\end{aligned}$$

Let $r(\mathcal{S}) = \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n$. Then, the above equation can be simplified as

$$\begin{aligned}\Delta_2(j \mid \mathcal{S}) &= \hat{F}_2(\mathcal{S} \cup \{j\}) - \hat{F}_2(\mathcal{S}) \\ &= \underbrace{\sum_{x \in Q(\mathcal{S}\cup\{j\})\setminus Q(\mathcal{S})} f(x) r(\mathcal{S}\cup\{j\})}_{T_1(\mathcal{S})} + \underbrace{\sum_{x\in Q(\mathcal{S})} f(x)\left(r(\mathcal{S}\cup\{j\}) - r(\mathcal{S})\right)}_{T_2(\mathcal{S})}\end{aligned}$$

It is easy to verify that $T_1(\mathcal{S})$ is monotone increasing function of $\mathcal{S}$. Let us consider $\mathcal{S}'$ such that $\mathcal{S} \subseteq \mathcal{S}' \subseteq \mathcal{C}\setminus\{j\}$. We have

$$\begin{aligned}\Delta_2(j \mid \mathcal{S}') - \Delta_2(j \mid \mathcal{S}) &\geq T_2(\mathcal{S}') - T_2(\mathcal{S}) \\ &\overset{T_2 \geq 0}{\geq} -g(\mathcal{S})(r(\mathcal{S}\cup\{j\}) - r(\mathcal{S}))\end{aligned}$$

Therefore, it suffices to find a lower bound on $-g(\mathcal{S})(r(\mathcal{S}\cup\{j\}) - r(\mathcal{S}))$. Further notice that

$$0 \leq g(\mathcal{S}) \leq \max_{\mathcal{T}:|\mathcal{T}|\leq |Q(\mathcal{S})|} \sum_{x\in\mathcal{T}} f(x) \tag{A.3}$$

and it is not hard to verify that

$$0 \leq r(\mathcal{S}\cup\{j\}) - r(\mathcal{S}) \leq \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n - \left(1 - \frac{1}{2|Q(\mathcal{S})|}\right)^n \tag{A.4}$$

Therefore, combining term (A.3) with (A.4), we get a lower bound on β:

$$\beta \geq -\max_{s\in\{1,\ldots,|Q(\mathcal{C})|\}} \left(\left(\left(1-\frac{1}{s}\right)^n - \left(1 - \frac{1}{2s}\right)^n\right) \underbrace{\max_{\mathcal{T}:|\mathcal{T}|\leq s} \sum_{x\in\mathcal{T}} f(x)}_{\text{Term 2}}\right) \tag{A.5}$$

Note that term 2 is a modular function and can be optimized greedily. Therefore, computing the RHS of Eq. A.5 can be efficiently done in polynomial time w.r.t. $|Q(\mathcal{C})|$. □

A.2 Proof of Lemma 3: Difference of Convex Construction of DS Decomposition

Lemma 5. *Let $g : 2^{\mathcal{C}} \to \mathbb{R}_{\geq 0}$ be a non-negative, non-decreasing supermodular function, and $u : \mathbb{R} \to \mathbb{R}$ be a non-decreasing convex function. For $\mathcal{S} \subseteq \mathcal{C}$, define $h(\mathcal{S}) = g(\mathcal{S}) \cdot u(|\mathcal{S}|)$. Then h is supermodular.*

Proof. Let $j \in \mathcal{C}$ and $\mathcal{S} \subseteq \mathcal{C} \setminus \{j\}$. The gain of j is

$$\begin{aligned}\Delta_h(j \mid \mathcal{S}) &= h(\mathcal{S}\cup\{j\}) - h(\mathcal{S}) \\ &= g(\mathcal{S}\cup\{j\}) \cdot u(|\mathcal{S}\cup\{j\}|) - g(\mathcal{S}) \cdot u(|\mathcal{S}|) \\ &= (g(\mathcal{S}\cup\{j\}) - g(\mathcal{S})) \cdot u(|\mathcal{S}\cup\{j\}|) + g(\mathcal{S})\left(u(|\mathcal{S}\cup\{j\}|) - u(|\mathcal{S}|)\right)\end{aligned}$$

Let us consider $\mathcal{S}'$ such that $\mathcal{S} \subseteq \mathcal{S}' \subseteq \mathcal{C} \setminus \{j\}$. We have

$$\begin{aligned}
\Delta_h(j \mid \mathcal{S}) &= (g(\mathcal{S} \cup \{j\}) - g(\mathcal{S})) \cdot u(|\mathcal{S} \cup \{j\}|) + g(\mathcal{S})\left(u(|\mathcal{S} \cup \{j\}|) - u(|\mathcal{S}|)\right) \\
&\overset{(a)}{\leq} (g(\mathcal{S}' \cup \{j\}) - g(\mathcal{S}')) \cdot u(|\mathcal{S}' \cup \{j\}|) + g(\mathcal{S})\left(u(|\mathcal{S} \cup \{j\}|) - u(|\mathcal{S}|)\right) \\
&\overset{(b)}{\leq} (g(\mathcal{S}' \cup \{j\}) - g(\mathcal{S}')) \cdot u(|\mathcal{S}' \cup \{j\}|) + g(\mathcal{S}')\left(u(|\mathcal{S}' \cup \{j\}|) - u(|\mathcal{S}'|)\right) \\
&= \Delta_h(j \mid \mathcal{S}')
\end{aligned}$$

where step (a) is due to g being monotone supermodular (i.e., $g(\mathcal{S}' \cup \{j\}) - g(\mathcal{S}') \geq g(\mathcal{S} \cup \{j\}) - g(\mathcal{S}) \geq 0$) and u being monotone (i.e., $u(|\mathcal{S}' \cup \{j\}|) \geq u(|\mathcal{S} \cup \{j\}|)$); step (b) is due to g being non-negative monotone (i.e., $g(\mathcal{S}') \geq g(\mathcal{S}) \geq 0$) and u being convex (i.e., $u(|\mathcal{S}' \cup \{j\}|) - u(|\mathcal{S}'|) \geq u(|\mathcal{S} \cup \{j\}|) - u(|\mathcal{S}|)$). Therefore h is supermodular. □

Lemma 6. *Let $w : \mathbb{R} \to \mathbb{R}$ be a convex function and $u : \mathbb{R} \to \mathbb{R}$ a convex non-decreasing function, then $u \circ w$ is convex. Furthermore, if w is non-decreasing, then the composition is also non-decreasing.*

Proof. By convexity of w:

$$w(\alpha x + (1 - \alpha)y) \leq \alpha w(x) + (1 - \alpha)w(y).$$

Therefore, we get

$$\begin{aligned}
u(w(\alpha x + (1 - \alpha)y)) &\overset{(a)}{\leq} u\left(\alpha w(x) + (1 - \alpha)w(y)\right) \\
&\overset{(b)}{\leq} \alpha u(w(x)) + (1 - \alpha)u(w(y)).
\end{aligned}$$

Here, step (a) is due to the fact that u is non-decreasing, and step (b) is due to the convexity of u. Therefore $u \circ w$ is convex. If w is non-decreasing, it is clear that $u \circ w$ is also non-decreasing, hence completes the proof. □

Lemma 7 (Horst & Thoai (1999))**.** *Let $r : \mathbb{R} \to \mathbb{R}$ be a non-decreasing, twice continuously differentiable function. Then r can be represented as the difference between two non-decreasing convex functions.*

Proof. Let $u : \mathbb{R} \to \mathbb{R}$ be a non-decreasing, strictly convex function, and $\alpha = \min_x u''(x)$; clearly, $\alpha > 0$. Let $\beta = |\min_x r''(x)|$. Define

$$v(x) = r(x) + \frac{\beta}{\alpha} u(x) \tag{A.6}$$

It is easy to verify that

$$v''(x) = r''(x) + \frac{\beta}{\alpha} u''(x) \geq r''(x) + \beta \geq 0.$$

Hence, $v(x)$ is convex. Furthermore, since both r and u are non-decreasing, v is also non-decreasing. Therefore, $r(x) = v(x) - \frac{\beta}{\alpha} u(x)$ is the difference between two non-decreasing convex functions. □

Lemma 8. *Let $r : \mathbb{R} \to \mathbb{R}$ be a non-decreasing, twice continuously differentiable function, and $w : \mathbb{R} \to \mathbb{R}$ a convex non-decreasing function, then $r \circ w$ can be represented as the difference between two non-decreasing convex functions.*

Proof. By Lemma 7, we can represent $r(x) = v(x) - \frac{\beta}{\alpha} u(x)$, where u, v are non-decreasing convex functions, and α, β are as defined in Eq. (A.6). Therefore,

$$r \circ w(x) = v \circ w(x) - \frac{\beta}{\alpha} \cdot u \circ w(x)$$

By Lemma 6, $v \circ w$ and $u \circ w$ are both non-decreasing convex, which completes the proof. □

Now we are ready to prove Lemma 3.

Proof of Lemma 3. Let $g(\mathcal{S}) = \sum_{x \in Q(\mathcal{S})} f(x)$. By definition we have

$$\hat{F}(\mathcal{S}) = g(\mathcal{S}) \left(1 - \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n\right) = g(\mathcal{S}) - g(\mathcal{S}) \left(1 - \frac{1}{|Q(\mathcal{S})|}\right)^n$$

Let $r(x) = \left(1 - \frac{1}{x}\right)^n$, and $w : \mathbb{R} \to \mathbb{R}$ be a convex function, such that $w(|\mathcal{S}|) = |Q(\mathcal{S})|$. Note that such function w exists, because the set function $h(\mathcal{S}) := |Q(\mathcal{S})|$ is supermodular. Therefore, we have

$$\hat{F}(\mathcal{S}) = g(\mathcal{S}) - g(\mathcal{S}) \cdot r \circ w(|\mathcal{S}|)$$

Furthermore, note that r is non-decreasing, twice continuously differentiable at $[1, \infty)$. By Lemma 8, we get

$$\begin{aligned} \hat{F}(\mathcal{S}) &= g(\mathcal{S}) - g(\mathcal{S}) \cdot \left(v \circ w(|\mathcal{S}|) - \frac{\beta}{\alpha} \cdot u \circ w(|\mathcal{S}|)\right) \\ &= g(\mathcal{S}) \left(1 + \frac{\beta}{\alpha} \cdot u \circ w(|\mathcal{S}|)\right) - g(\mathcal{S}) \cdot (v \circ w(|\mathcal{S}|)), \end{aligned} \tag{A.7}$$

where $u : \mathbb{R} \to \mathbb{R}$ can be any non-decreasing, strictly convex function, $\alpha = \min_x u''(x)$, $\beta = |\min_{x \geq 1} r''(x)|$, and $v(x) = r(x) + \frac{\beta}{\alpha} u(x)$.

We know from Lemma 4 that g is supermodular. Since both $1 + \frac{\beta}{\alpha} \cdot u \circ w(x)$ and $v \circ w(x)$ are convex, then by Lemma 5, we know that both terms on the R.H.S. of Eq. (A.7) are supermodular, and hence we obtain a DS decomposition of function $\hat{F}$. $\square$

B Supplemental Figures

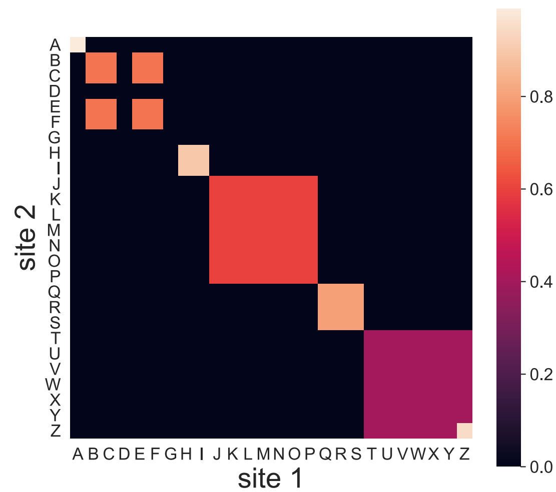
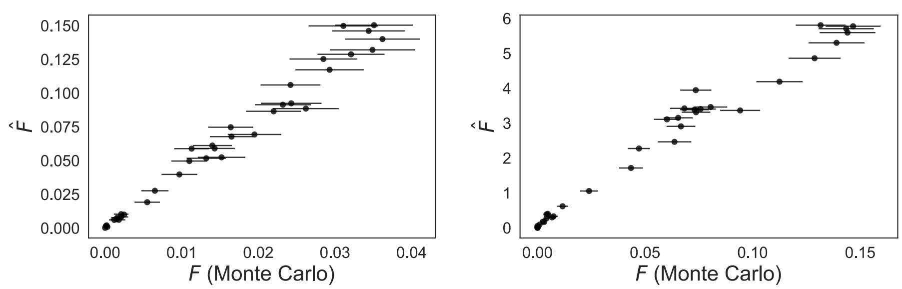


Figure S1: The cell values for the synthetic dataset with $L = 2$ and $|\mathcal{C}^{(\ell)}| = 26\ \forall \ell \in \{1, 2\}$.



(a) $\hat{F}$ vs. F for GB1

(b) $\hat{F}$ vs. F for PhoQ

Figure S2: Comparing $\hat{F}$ (Eq. (3.2)) against the Monte Carlo estimates of F (Eq. (3.1)). Error bars are standard errors for the Monte Carlo estimates. The approximate objective correlates well with Monte Carlo estimates of the exact objective.