A Proof of Theorem 1

Because P is a bistochastic matrix, and we know $P^* = \frac{1}{\sqrt{n}}\mathbb{1}$, we can lower bound

$$\begin{aligned} 1 \geq P(w)^T P^* &= \sum_i w_i \frac{P_i^T}{\|P_i\|} \frac{1}{\sqrt{n}}\mathbb{1} \\ &= \frac{1}{\sqrt{n}} \sum_i \frac{w_i}{\|P_i\|} \\ &\geq \frac{1}{\min \|P_i\| \sqrt{n}} \end{aligned}$$

Similarly, $C(S) \leq C^k_{max}$. Now given $\lambda \leq \frac{1-\kappa}{C^k_{max} \min \|P_i\| \sqrt{n}}$, we compute

$$\begin{aligned} \max_w P(w)^T P^* - \lambda C(S) &\geq \max_w P(w)^T P^* - \frac{1-\kappa}{\min \|P_i\| \sqrt{n}} \\ &\geq \max_w P(w)^T P^* \left(1 - \frac{1-\kappa}{\min \|P_i\| \sqrt{n}} \frac{1}{P(w)^T P^*}\right) \\ &\geq max_w P(w)^T P^* \left(1 - \frac{1-\kappa}{\min \|P_i\| \sqrt{n}} \min \|P_i\| \sqrt{n}\right) \\ &\geq \kappa \max_w P(w)^T P^*. \end{aligned}$$

B Auxiliary Lemmas

First, we make a few statements related to initialization of the process. Lemma 3.4 from [6] directly applies to this problem, and thus $\delta_k \in [0, 1]\ \forall k$.

Lemma 4. $\langle Pw_1, \frac{1}{n}\mathbb{1}\rangle \geq \frac{\kappa}{\sqrt{n}\sum_i \|P_i\|}$

Proof follows equivalently to Lemma 3.1 from [6], with added caveat that our choice of weights is within κ of maximum value.

Lemma 5. *The cost aware geodesic alignment $\langle a_t, a_{t,v_k}\rangle$ satisfies*

$$\langle a_k, a_{k,v_k}\rangle \geq \kappa\tau\sqrt{J_t} \vee f(t),$$

for

$$f(x) = \kappa \frac{\sqrt{1-x}\sqrt{1-\beta^2}\epsilon + \sqrt{x}\beta}{\sqrt{1 - \left(\sqrt{x}\sqrt{1-\beta^2}\epsilon - \sqrt{1-x}\beta\right)^2}}$$

and

$$\beta = 0 \wedge \min\langle \ell_n, \frac{1}{\sqrt{n}}\mathbb{1}\rangle \text{ s.t. } \langle \ell_n, \frac{1}{\sqrt{n}}\mathbb{1}\rangle > -1.$$

Proof. The lemma is equivalent to proving Lemma 3.6 in [6] with one caveat. Here our choice of node is v_k, which comes from choosing the cheapest cost node location from the set $S = \{v \in V | \langle a_k, a_{kv}\rangle \geq \kappa \langle a_k, a_{kv_k}\rangle\}$ Because of this, we can recover all results from $\langle a_k, a_{kv_k}\rangle$ with only a constant κ in front, as our choice satisfies $\langle a_k, a_{kv_k}\rangle \geq \kappa \langle a_k, a_{kv_k}\rangle$. □

We apply Lemma 5 to prove the following Theorem that is needed, and mirrors the results from [6].

Theorem 6. *Assume a cost of sensor placement $C(v) : V \to \mathbb{R}_+$ and a slack parameter κ. If we choose the set of points W and weights a_w using Algorithm 1 such that $|W| = K$, then*

$$\|Pw - \frac{1}{n}\mathbb{1}\| \leq \frac{\eta\, v_K}{\sqrt{n}},$$

where $v_K = O((1-\kappa^2\epsilon^2)^{K/2})$ for some ϵ and $\eta = \sqrt{1 - \kappa^2 \max_{i\in V} \left\langle \frac{P_i}{\|P_i\|}, \frac{1}{\sqrt{n}}\mathbb{1}\right\rangle^2}$.

Proof. We mimic the results from [6], incorporating the additional cost parameter. We denote $J_k := 1 - \langle \frac{Pw_k}{\|Pw_k\|}, \frac{1}{\sqrt{n}}\mathbb{1}\rangle$. If we substitute this into the formula for δ_t, we get

$$J_{k+1} = J_k(1 - \langle a_t, a_{kv_k}\rangle^2).$$

Applying our bound from Lemma 5, we get

$$J_{k+1} \leq J_k \left(1 - \kappa^2\tau^2 J_k\right).$$

By applying the standard induction argument used in [6], we get

$$J_k \leq B(k) := \frac{J_1}{1 + \kappa^2\tau^2(k-1)}.$$

Because $B(k)$ still goes to 0, and $f(B(k)) \to \kappa\epsilon$, there exists a k^* such that $f(B(k)) \geq \kappa\tau\sqrt{B(k)}$, and since f is monotonic decreasing, $f(J_t) > f(B(k))$. Using Lemma 5, we finish with

$$J_k \leq B(k \wedge k^*) \prod_{s=k^*+1}^{k} (1 - f^2(B(s)))$$

We note that $\frac{1}{n}J_k = \|\beta^* P(w) - P^*\|^2$, so this means

$$\|\beta^* Pw - \frac{1}{n}\mathbb{1}\| \leq \frac{\eta\, C_K}{\sqrt{n}},$$

for constant C_K combining the denominator in $B(k)$ and the product of $\prod_{s=k^*+1} k1 - f^2(B(s))$, and $\sqrt{J_1} = \eta$. Notice that $f(B(k)) \to \kappa\epsilon$ shows a rate of decay of $v = \sqrt{1 - \kappa^2\epsilon^2}$. □

C Proof of Theorem 2

We note that [27] proves multiple bounds on $\left|\frac{1}{n}\sum_{v\in V} f(v) - \sum_{s\in S} w_s f(s)\right|$. The main bound in the paper comes from using the fact that they assume $\sum_s w_s = 1$, which allows them to break up the inner product $\|P\sum_s w_s\delta_s - \frac{1}{n}\mathbb{1}\|$ into its subsequent terms $\left(\|P\sum_s w_s\delta_w\|^2 - \frac{1}{n}\right)^{1/2}$. We step away from this assumption and will instead work directly with the norm $\left\|P\sum_s w_s\delta_s - \frac{1}{n}\mathbb{1}\right\| = \left\|Pw - \frac{1}{n}\mathbb{1}\right\|$.

By the same logic as in [27], we know

$$\left|\frac{1}{n}\sum_{v\in V} f(v) - \sum_{s\in S} w_s f(s)\right| \leq \frac{\|f\|_{P_\lambda}}{\lambda^\ell} \min_{\beta, w} \|\beta P^\ell w - \frac{1}{n}\mathbb{1}\|.$$

We can simply replace $\widetilde{P} = P^\ell$ and inherit on $\widetilde{P}$ in Theorem 6, in particular that we still have $\sum_i \frac{1}{n}\widetilde{P}_i = \frac{1}{n}\mathbb{1}$. Thus, we can apply the guarantees of Algorithm 1 and Theorem 6 to bound $\|\beta P^\ell w - \frac{1}{n}\mathbb{1}\| \leq \frac{\eta\, v_K}{\sqrt{n}}$ and attain the desired result.