# SUPPLEMENTARY MATERIAL FOR TROPICAL SUPPORT VECTOR MACHINE AND ITS APPLICATIONS TO PHYLOGENOMICS

BY XIAOXIAN TANG[1], HOUJIE WANG[2] AND RURIKO YOSHIDA[3]

[1]*School of Mathematical Sciences, Beihang University, xiaoxian@buaa.edu.cn*

[2]*Department of Statistics, Texas A&M University, wanghoujie@tamu.edu*

[3]*Department of Operations Research, Naval Postgraduate School, ryoshida@nps.edu*

This is a supplementary file for the main manuscript entitled "Tropical Support Vector Machine and its Applications to Phylogenomics". It is organized as follows: In Section 1, we provide the information on all files (including supplementary material, code and data) at the online repository. In Section 2, we give proofs of all theorems in the main manuscript. In Section 3, we present pseudo code of algorithms for computing soft margin tropical SVMs, which are implemented and applied to simulated data sets in the main manuscript. Also, we explain these algorithms in details.

**1. Supplementary Information.** Table 1 lists all files at the online repository: `https://github.com/HoujieWang/Tropical-SVM`

TABLE 1
*Supporting Information files*

| Name | File Type | Results |
|---|---|---|
| `Supplementary_Material.pdf` | `PDF` | Information table, proofs and algorithms |
| `unbounded.RData` | `RData` | Remark for Theorem 4.6 |
| `Algorithm1.R` | `R` | Algorithm 1 |
| `Algorithm2.R` | `R` | Algorithm 2 |
| `Algorithm3.R` | `R` | Algorithm 3 |
| `Algorithm4.R` | `R` | Algorithm 4 |
| `graph producer.R` | `R` | Figure 5 |
| `Data` | folder | Figure 5 |
| `Data/data_15%.RData` | `R Data` | Figure 5 |
| `Data/data_20%.RData` | `R Data` | Figure 5 |
| `Data/data_25%.RData` | `R Data` | Figure 5 |
| `Data/asgn_1_15%.RData` | `R Data` | Figure 5 |
| `Data/asgn_2_15%.RData` | `R Data` | Figure 5 |
| `Data/asgn_3_15%.RData` | `R Data` | Figure 5 |
| `Data/asgn_4_15%.RData` | `R Data` | Figure 5 |
| `Data/asgn_1_20%.RData` | `R Data` | Figure 5 |
| `Data/asgn_2_20%.RData` | `R Data` | Figure 5 |
| `Data/asgn_3_20%.RData` | `R Data` | Figure 5 |
| `Data/asgn_4_20%.RData` | `R Data` | Figure 5 |
| `Data/asgn_1_25%.RData` | `R Data` | Figure 5 |
| `Data/asgn_2_25%.RData` | `R Data` | Figure 5 |
| `Data/asgn_3_25%.RData` | `R Data` | Figure 5 |
| `Data/asgn_4_25%.RData` | `R Data` | Figure 5 |

**2. Proofs.** In this section, we prove all theorems in the main manuscript. We first recall formulations for tropical hard margin SVMs and tropical soft margin SVMs. Given two finite and tropically separable sets $P$ and $Q$ in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$, assume $|P|=|Q|=n>0$ (remark that all our results can be directly extended when $|P|\neq|Q|$). For any $\xi\in P\cup Q$, we denote by $i(\xi)$ and $j(\xi)$ two indices in terms of $\xi$, which are integers in the set $\{1,\ldots,d\}$. We denote the two sets of indices $\{i(\xi)|\xi\in P\cup Q\}$ and $\{j(\xi)|\xi\in P\cup Q\}$ by $\mathcal{I}$ and $\mathcal{J}$, respectively. We also assume that

$$\forall\xi\in P\cup Q, i(\xi)\neq j(\xi), \quad \text{and} \tag{1}$$

$$\forall p\in P,\ \forall q\in Q,\ \ i(p)\neq i(q). \tag{2}$$

We first recall a tropical hard margin SVM for solving the normal vector $\omega$ of an optimal tropical separating hyperplane $H_\omega$ for $P$ and $Q$:

$$\max_{z\in\mathbb{R}} z \tag{3}$$

$$\text{s.t. } \forall\xi\in P\cup Q,\ \ z+\xi_{j(\xi)}+\omega_{j(\xi)}-\xi_{i(\xi)}-\omega_{i(\xi)}\leq 0, \tag{4}$$

$$\forall\xi\in P\cup Q,\ \omega_{j(\xi)}-\omega_{i(\xi)}\leq\xi_{i(\xi)}-\xi_{j(\xi)}, \tag{5}$$

$$\forall\xi\in P\cup Q,\forall l\neq i(\xi),j(\xi),\ \ \omega_l-\omega_{j(\xi)}\leq\xi_{j(\xi)}-\xi_l. \tag{6}$$

Below, we recall a tropical soft margin SVM:

$$\max_{(z;\alpha;\beta;\gamma)\in\mathbb{R}^{2dn+1}} z-\mathcal{C}\sum_{\xi\in P\cup Q}\left(\alpha_\xi+\beta_\xi+\sum_{l\neq i(\xi),j(\xi)}\gamma_{\xi,l}\right) \tag{7}$$

$$\text{s.t. } \forall\xi\in P\cup Q,\ \ z+\xi_{j(\xi)}+\omega_{j(\xi)}-\xi_{i(\xi)}-\omega_{i(\xi)}\leq\alpha_\xi, \tag{8}$$

$$\forall\xi\in P\cup Q,\ \ \omega_{j(\xi)}-\omega_{i(\xi)}\leq\xi_{i(\xi)}-\xi_{j(\xi)}+\beta_\xi, \tag{9}$$

$$\forall\xi\in P\cup Q,\ \forall l\neq i(\xi),j(\xi),\ \ \omega_l-\omega_{j(\xi)}\leq\xi_{j(\xi)}-\xi_l+\gamma_{\xi,l}, \tag{10}$$

$$\forall\xi\in P\cup Q,\ \forall l\neq i(\xi),j(\xi),\ \ \alpha_\xi\geq 0,\ \ \beta_\xi\geq 0,\ \ \gamma_{\xi,l}\geq 0,\ \ \text{and } z\geq 0, \tag{11}$$

2.1. *Proofs in Section 4.1 (hard margin tropical SVMs).* Theorems 2.2–2.6 below are respectively Theorems 4.1–4.5 in the main manuscript.

DEFINITION 2.1 (Tropically Separable Sets and Tropical Separating Hyperplane). *For any two finite sets $P$ and $Q$ in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$, if there exists $\omega\in\mathbb{R}^d$ such that*

(i) *for any $\xi\in P\cup Q$, there exists an index $i\in\{1,\ldots,d\}$ such that*

$$\textit{for any } j\in\{1,\ldots,d\}\backslash\{i\},\ \ \omega_i+\xi_i\ >\ \omega_j+\xi_j,\ \ \textit{and}$$

(ii) *for any $p\in P$, and for any $q\in Q$, we have*

$$\operatorname*{argmax}_{1\leq k\leq d}\{\omega_k+p_k\}\ \neq\ \operatorname*{argmax}_{1\leq k\leq d}\{\omega_k+q_k\},$$

*then we say $P$ and $Q$ are* tropically separable*, and we say $H_\omega$ is a* tropical separating hyperplane *for $P$ and $Q$.*

THEOREM 2.2. *If two finite sets $P$ and $Q$ in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$ are tropically separable, then there exist $\omega\in\mathbb{R}^d$ and two sets of indices $\mathcal{I}:=\{i(\xi)|\xi\in P\cup Q\}$ and $\mathcal{J}:=\{j(\xi)|\xi\in P\cup Q\}$ such that the constraints (1)–(2) and (4)–(6) are satisfied. More than that, if $(z^*;\omega^*)$ is an optimal solution to (3)–(6) w.r.t. $\mathcal{I}$ and $\mathcal{J}$, then the margin $z^*$ (i.e., the optimal value of (3)) is positive, and $H_{\omega^*}$ is a separating tropical hyperplane for $P$ and $Q$.*

PROOF. In fact, by Definition 2.1, there exist indices $\mathcal{I}$ and $\mathcal{J}$ and a tropical separating hyperplane $H_\omega$ such that (1)–(2) and (4)–(6) are satisfied. The condition (i) in Definition 2.1 ensures that the distance $d_{\rm tr}(\xi, H_\omega)$ is nonzero for any $\xi \in P \cup Q$, and hence, the value of $z$ corresponding to $\omega$ (i.e., the minimum distance from the points in $P \cup Q$ to $H_\omega$) is positive. If $(z^*;\omega^*)$ is an optimal solution w.r.t. $\mathcal{I}$ and $\mathcal{J}$, then $z^* \geq z > 0$. Note $(z^*;\omega^*)$ is also a feasible solution, so (4)–(6) also hold for $\omega^*$ w.r.t. $\mathcal{I}$ and $\mathcal{J}$. So, the condition (i) in Definition 2.1 holds for $\omega^*$. By (2), any two points $p$ and $q$ from different sets will be located in different open sectors $S_{\omega^*}^{i(p)}$ and $S_{\omega^*}^{i(q)}$. So, the the condition (ii) in Definition 2.1 is also satisfied for $\omega^*$. Therefore, $H_{\omega^*}$ is a separating tropical hyperplane for $P$ and $Q$. □

THEOREM 2.3. *Suppose $P$ and $Q$ are two finite sets in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. For all $p \in P$, assume $i(p)$ and $j(p)$ are constants, say $i_P$ and $j_P$. For all $q \in Q$, assume $i(q)$ and $j(q)$ are constants, say $i_Q$ and $j_Q$. If the four numbers $i_P, j_P, i_Q$ and $j_Q$ are pairwise distinct, then the linear programming (3)–(6) has a feasible solution if and only if*

$$\max\{-F,\, -A-E\} \,\leq\, \min\{D+B,\, C\}. \tag{12}$$

*If a feasible solution exists, then the optimal value $z$ is given by*

$$\min\left\{\, A+C+E,\; D+B+F,\; \frac{1}{2}(A+B+D+E)\,\right\}, \tag{13}$$

*where*

$$\begin{aligned} A &= \min_{p\in P}\{p_{i_P} - p_{j_P}\},\; C = \min_{p\in P}\{p_{j_P} - p_{j_Q}\},\; E = \min_{q\in Q}\{q_{j_Q} - q_{i_P}\},\\ B &= \min_{p\in P}\{p_{j_P} - p_{i_Q}\},\; D = \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\},\; F = \min_{q\in Q}\{q_{j_Q} - q_{j_P}\}. \end{aligned} \tag{14}$$

PROOF. For any $p \in P$, by the assumptions $i(p) = i_P$ and $j(p) = j_P$, and by (5)–(6), we have:

$$\omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P}, \text{ and } \omega_l - \omega_{j_P} \leq p_{j_P} - p_l,\;\; \forall l \neq i_P, j_P.$$

By the assumption that the four numbers $i_P, j_P, i_Q$ and $j_Q$ are pairwise distinct, there exists $l \neq i_P, j_P$ such that $l = i_Q$ or $l = j_Q$. Similarly, for any $q \in Q$, by (5)–(6), we have

$$\omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q}, \text{ and } \omega_l - \omega_{j_Q} \leq q_{j_Q} - q_l,\;\; \forall l \neq i_Q, j_Q,$$

and there exists $l \neq i_Q, j_Q$ such that $l = i_P$ or $l = j_P$. So for any $p \in P$, we have

$$\omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P}, \tag{15}$$

$$\omega_{i_Q} - \omega_{j_P} \leq p_{j_P} - p_{i_Q}, \tag{16}$$

$$\omega_{j_Q} - \omega_{j_P} \leq p_{j_P} - p_{j_Q}. \tag{17}$$

and for any $q \in Q$,

$$\omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q}, \tag{18}$$

$$\omega_{i_P} - \omega_{j_Q} \leq q_{j_Q} - q_{i_P}, \tag{19}$$

$$\omega_{j_P} - \omega_{j_Q} \leq q_{j_Q} - q_{j_P}. \tag{20}$$

By adding (16) and (18), we have

$$\omega_{j_Q} - \omega_{j_P} \,\leq\, \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\}. \tag{21}$$

By adding (15) and (19), we have

$$\max_{p\in P}\{p_{j_P} - p_{i_P}\} + \max_{q\in Q}\{q_{i_P} - q_{j_Q}\} \,\leq\, \omega_{j_Q} - \omega_{j_P}. \tag{22}$$

By (17), (20), (21) and (22), the inequality (12) holds. Therefore, if the linear programming (3)–(6) has a feasible solution, then we have (12).

On the other hand, if we have (12), then there exist real numbers $\omega_{j_Q}$ and $\omega_{j_P}$ such that the inequalities (17), (20), (21) and (22) hold. Notice that the inequality (21) is equivalent to

$$\omega_{j_Q} - \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} \le \omega_{j_P} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\}.$$

So, there exists a number $\omega_{i_Q}$ such that

$$(23) \qquad \omega_{j_Q} - \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} \le \omega_{i_Q}, \qquad\qquad (24) \qquad \omega_{i_Q} \le \omega_{j_P} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\}.$$

Symmetrically, the inequality (22) is equivalent to

$$\max_{p\in P}\{p_{j_P} - p_{i_P}\} + \omega_{j_P} \le \omega_{j_Q} - \max_{q\in Q}\{q_{i_P} - q_{j_Q}\}.$$

So, there exists a number $\omega_{i_P}$ such that

$$(25) \qquad \max_{p\in P}\{p_{j_P} - p_{i_P}\} + \omega_{j_P} \le \omega_{i_P}, \qquad\qquad (26) \qquad \omega_{i_P} \le \omega_{j_Q} - \max_{q\in Q}\{q_{i_P} - q_{j_Q}\}.$$

The inequality (25) can be rewritten as

$$(27) \qquad \max_{p\in P}\{p_{j_P} - p_{i_P}\} \le \omega_{i_P} - \omega_{j_P} \Leftrightarrow \omega_{j_P} - \omega_{i_P} \le \min_{p\in P}\{p_{i_P} - p_{j_P}\}.$$

The inequality (26) can be rewritten as

$$(28) \qquad \omega_{i_P} - \omega_{j_Q} \le -\max_{q\in Q}\{q_{i_P} - q_{j_Q}\} \Leftrightarrow \omega_{i_P} - \omega_{j_Q} \le \min_{q\in Q}\{q_{j_Q} - q_{i_P}\}.$$

By (23) and (27), the inequality (5) holds. By (17), (20), (24) and (28), the inequality (6) holds for $l = i_Q, j_Q$ when $\xi \in P$, or for $l = i_P, j_P$ when $\xi \in Q$. For $l \ne i_Q, j_Q$ when $\xi \in P$, or for $l \ne i_P, j_P$ when $\xi \in Q$, there always exist sufficiently small numbers for $\omega_l$ such that the inequality (6) holds. So the inequality (12) guarantees the feasibility of the inequalities (5) and (6). Notice that once (5) and (6) are feasible, there is always a non-negative number $z$ such that (4) holds. So, if we have (12), then the linear programming (3)–(6) has a feasible solution.

If a feasible solution exists, then by (4), for any feasible solution $(\omega; z)$,

$$(29) \qquad z \le \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \omega_{i_P} - \omega_{j_P}, \;\; z \le \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_Q}.$$

So, by (16) and (19), and by summing up the above two inequalities, we have

$$\begin{aligned}
(30) \qquad 2z &\le \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_P} - \omega_{j_P} + \omega_{i_Q} - \omega_{j_Q} \\
&= \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + (\omega_{i_Q} - \omega_{j_P}) + (\omega_{i_P} - \omega_{j_Q}) \\
(31) \qquad &\le \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\} + \min_{q\in Q}\{q_{j_Q} - q_{i_P}\} \\
&= A + D + B + E.
\end{aligned}$$

Also, by (17), (19) and the first inequality in (29), we have

$$(32) \qquad z \le \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \omega_{i_P} - \omega_{j_P}$$

$$= \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \omega_{i_P} - \omega_{j_Q} + \omega_{j_Q} - \omega_{j_P}$$

$$= \min_{p\in P}\{p_{i_P} - p_{j_P}\} + (\omega_{j_Q} - \omega_{j_P}) + (\omega_{i_P} - \omega_{j_Q})$$

$$\le \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{p\in P}\{p_{j_P} - p_{j_Q}\} + \min_{q\in Q}\{q_{j_Q} - q_{i_P}\} \;=\; A + C + E. \tag{33}$$

Symmetrically, by (16), (20) and the second inequality in (29), we have

$$z \le \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_Q} \tag{34}$$

$$= \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + (\omega_{i_Q} - \omega_{j_P}) + (\omega_{j_P} - \omega_{j_Q})$$

$$\le \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\} + \min_{q\in Q}\{q_{j_Q} - q_{j_P}\} \;=\; D + B + F. \tag{35}$$

Hence, all the values $\frac{1}{2}(A+B+D+E)$, $A+C+E$ and $B+D+F$ are upper bounds for $z$.

If $\frac{1}{2}(A+B+D+E) \le \min\{A+C+E, B+D+F\}$, then we can choose a feasible $\omega$ such that

$$\omega_{i_Q} - \omega_{j_P} = -B,\ \omega_{i_P} - \omega_{j_Q} = -E,$$

$$\omega_{j_P} - \omega_{i_P} = \frac{1}{2}(A-B-D-E),\ \omega_{j_Q} - \omega_{i_Q} = \frac{1}{2}(D-A-B-E)$$

which imply $\omega_{j_Q} - \omega_{j_P} = \frac{1}{2}(B+D-E-A)$. For this $\omega$, the equalities in (29), (30) and (31) hold, and $z$ reaches its optimal value $\frac{1}{2}(A+B+D+E)$.

If $A+C+E < \min\{\frac{1}{2}(A+B+D+E),\ B+D+F\}$, then we can choose a feasible $\omega$ such that

$$\omega_{j_Q} - \omega_{j_P} = C, \text{ and } \omega_{i_P} - \omega_{j_Q} = E,$$

which imply $\omega_{j_P} - \omega_{i_P} = -C-E$. For this $\omega$, the equalities in (32) and (33) hold, and $z$ reaches its optimal value $A+C+E$. Symmetrically, if

$$B+D+F < \min\{\frac{1}{2}(A+B+D+E),\ A+C+E\},$$

then we can choose a feasible $\omega$ such that

$$\omega_{j_P} - \omega_{j_Q} = F, \text{ and } \omega_{i_Q} - \omega_{j_P} = B,$$

which imply $\omega_{j_Q} - \omega_{i_Q} = -B-F$. For this $\omega$, the equalities in (34) and (35) hold, and $z$ reaches its optimal value $B+D+F$. □

THEOREM 2.4. *Suppose $P$ and $Q$ are two finite sets in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. For all $p \in P$, assume $i(p)$ and $j(p)$ are constants, say $i_P$ and $j_P$. For all $q \in Q$, assume $i(q)$ and $j(p)$ are constants, say $i_Q$ and $j_Q$.*

(i) *If $i_P = j_Q$ and $i_Q \ne j_P$, then linear programming (3)–(6) has a feasible solution if and only if*

$$A + B + C \ge 0. \tag{36}$$

*If a feasible solution exists, then the optimal value $z$ is given by*

$$\min\{\ A+B+C,\ \frac{1}{2}\big(A'+B+C\big)\},$$

*where*

$$A' = \min_{p\in P}\{p_{i_P} - p_{j_P}\},\ A = \min_{\xi\in P\cup Q}\{\xi_{i_P} - \xi_{j_P}\},$$
$$B = \min_{p\in P}\{p_{j_P} - p_{i_Q}\},\quad C = \min_{q\in Q}\{q_{i_Q} - q_{i_P}\}.$$

(ii) *If $i_Q = j_P$ and $i_P \neq j_Q$, then linear programming (3)–(6) has a feasible solution if and only if*

$$A + B + C \geq 0. \tag{37}$$

*If a feasible solution exists, then the optimal value $z$ is given by*

$$\min\{\ A + B + C,\ \frac{1}{2}\big(A' + B + C\big)\ \},$$

*where*

$$A' = \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\},\ A = \min_{\xi\in P\cup Q}\{\xi_{i_Q} - \xi_{j_Q}\},$$
$$B = \min_{q\in Q}\{q_{j_Q} - q_{i_P}\},\quad C = \min_{p\in P}\{p_{i_P} - p_{i_Q}\}.$$

PROOF. We only need to prove part (i) since part (ii) can be symmetrically argued. For any $p \in P$, by the assumptions $i(p) = i_P$ and $j(p) = j_P$, and by (5)–(6), we have:

$$\omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P}, \text{ and } \omega_l - \omega_{j_P} \leq p_{j_P} - p_l,\ \ \forall l \neq i_P, j_P.$$

By (1), $i_P \neq i_Q$. Note that we assume $i_Q \neq j_P$. So, there exists $l \neq i_P, j_P$ such that $l = i_Q$. For any $q \in Q$, by (5)–(6), we have

$$\omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q}, \text{ and } \omega_l - \omega_{j_Q} \leq q_{j_Q} - q_l,\ \ \forall l \neq i_Q, j_Q.$$

By the definition of $i(p)$ and $j(p)$, we have $i_P \neq j_P$. So we have $j_Q \neq j_P$ since we assume that $i_P = j_Q$. Notice again that we assume $i_Q \neq j_P$. Hence, there exists $l \neq i_Q, j_Q$ such that $l = j_P$. So, for any $p \in P$,

$$\omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P}, \tag{38}$$

$$\omega_{i_Q} - \omega_{j_P} \leq p_{j_P} - p_{i_Q}, \tag{39}$$

and for any $q \in Q$,

$$\omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q}, \tag{40}$$

$$\omega_{j_P} - \omega_{j_Q} \leq q_{j_Q} - q_{j_P}. \tag{41}$$

If $i_P = j_Q$, then by (38) and (41), we have

$$\max_{\xi\in P\cup Q}\{\xi_{j_P} - \xi_{i_P}\}\ \leq\ \omega_{i_P} - \omega_{j_P}. \tag{42}$$

If $i_P = j_Q$, then by adding (39) and (40), we have

$$\omega_{i_P} - \omega_{j_P}\ \leq\ \min_{p\in P}\{p_{j_P} - p_{i_Q}\} + \min_{q\in Q}\{q_{i_Q} - q_{i_P}\}. \tag{43}$$

So, if $i_P = j_Q$, then by (42) and (43), we have (36).

On the other hand, if we have (36), then there exist real numbers $\omega_{i_P}$ and $\omega_{j_P}$ such that (42) and (43) hold. By (42), we have (38) and (41). Let $\omega_{i_Q} = \omega_{i_P} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\}$. Then we have the inequality (39), and by (43), we have

$$\max_{q\in Q}\{q_{i_P} - q_{i_Q}\}\ \leq\ \omega_{i_Q} - \omega_{i_P},$$

which is equivalent to (40). By (38), (39), (40) and (41), the inequality (5) holds, and the the inequality (6) holds for $l = i_Q$ when $\xi \in P$, or for $l = j_P$ when $\xi \in Q$. For $l \neq i_Q$ when $\xi \in P$, or for $l \neq j_P$ when $\xi \in Q$, there always exist sufficiently small numbers for $\omega_l$ such that the inequality (6) holds. So the inequality (36) guarantees the feasibility of the inequalities (5) and (6). Notice that once (5) and (6) are feasible, there is always a non-negative number $z$ such that (4) holds. So, if we have (36), then the linear programming (3)–(6) has a feasible solution.

If a feasible solution exists, then by (4),

$$z \leq \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \omega_{i_P} - \omega_{j_P}, \ \ z \leq \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_Q}. \tag{44}$$

Note $i_P = j_Q$. So, by summing up the above two inequalities and by (39),

$$\begin{aligned} 2z \ &\leq \ \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_P} && (45)\\ &\leq \ \min_{p\in P}\{p_{i_P} - p_{j_P}\} + \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\} \ = \ A' + C + B. && (46)\end{aligned}$$

Also, by (39), (42) and the second inequality in (44), we have

$$\begin{aligned} z \ &\leq \ \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_Q} && (47)\\ &= \ \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \omega_{i_Q} - \omega_{j_P} + \omega_{j_P} - \omega_{j_Q}\\ &= \ \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + (\omega_{i_Q} - \omega_{j_P}) + (\omega_{j_P} - \omega_{i_P})\\ &\leq \ \min_{q\in Q}\{q_{i_Q} - q_{j_Q}\} + \min_{p\in P}\{p_{j_P} - p_{i_Q}\} + \min_{\xi\in P\cup Q}\{\xi_{i_P} - \xi_{j_P}\} \ = \ C + B + A. && (48)\end{aligned}$$

Hence, both values $\frac{1}{2}(A' + B + C)$ and $A + B + C$ are upper bounds for $z$.

If $\frac{1}{2}(A' + B + C) \leq A + B + C$, then we can choose a feasible $\omega$ such that

$$\omega_{i_P} - \omega_{j_P} = \frac{1}{2}\left(B + C - A'\right), \text{ and } \omega_{i_Q} = \omega_{j_P} + B,$$

which imply $\omega_{i_Q} - \omega_{j_Q} = \frac{1}{2}(A' + B - C)$. For this $\omega$, the equalities in (44)–(46) hold, and $z$ reaches its optimal value $\frac{1}{2}(A' + B + C)$.

If $A + B + C < \frac{1}{2}(A' + B + C)$, then we can choose a feasible $\omega$ such that

$$\omega_{i_Q} - \omega_{j_Q} = A + B, \text{ and } \omega_{i_Q} = \omega_{j_P} + B,$$

which imply $\omega_{i_P} - \omega_{j_P} = -A$. For this $\omega$, the equalities in (47)–(48) hold, and $z$ reaches its optimal value $A + B + C$. □

THEOREM 2.5. *Suppose $P$ and $Q$ are two finite sets in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. If for all $p \in P$, we have $i(p) = j(q) = k_1$, and for all $q \in Q$, $i(q) = j(p) = k_2$, then the linear programming (3)–(6) has a feasible solution if and only if*

$$\max_{p\in P}\{p_{k_2} - p_{k_1}\} \leq \min_{q\in Q}\{q_{k_2} - q_{k_1}\} \tag{49}$$

*If a feasible solution exists, then the optimal value $z$ is given by*

$$\frac{1}{2}(\min_{p\in P}\{p_{k_1} - p_{k_2}\} + \min_{q\in Q}\{q_{k_2} - q_{k_1}\}) \tag{50}$$

PROOF. For any $p \in P$, and for any $q \in Q$, by the assumptions $i(p) = j(q) = k_1$ and $i(q) = j(p) = k_2$, and by (5), we have:

$$\omega_{k_2} - \omega_{k_1} \le p_{k_1} - p_{k_2}, \text{ and } \omega_{k_1} - \omega_{k_2} \le q_{k_2} - q_{k_1}$$

Therefore,

$$\max_{p\in P}\{p_{k_2} - p_{k_1}\} \le \omega_{k_1} - \omega_{k_2} \le \min_{q\in Q}\{q_{k_2} - q_{k_1}\}. \tag{51}$$

So if the linear programming (3)–(6) has a feasible solution, then we have (49).

On the other hand, if we have (49), then there exist real numbers $\omega_{k_1}$ and $\omega_{k_2}$ such that (51), and hence (5) holds. For $l \ne k_1, k_2$, there always exist sufficiently small numbers for $\omega_l$ such that the inequality (6) holds. So inequalities (49) guarantees the feasibility of the inequalities (5) and (6). Notice that once (5) and (6) are feasible, there is always a non-negative number $z$ such that (4) holds. So, if we have (49), then the linear programming (3)–(6) has a feasible solution.

If a feasible solution exists, then by (4),

$$z \le \min_{p\in P}\{p_{k_1} - p_{k_2}\} + \omega_{k_1} - \omega_{k_2},\ \ z \le \min_{q\in Q}\{q_{k_2} - q_{k_1}\} + \omega_{k_2} - \omega_{k_1}. \tag{52}$$

By summing up the above two inequalities,

$$2z \le \min_{p\in P}\{p_{k_1} - p_{k_2}\} + \min_{q\in Q}\{q_{k_2} - q_{k_1}\}.$$

So the value (50) is an upper bound for $z$. Note that we can choose a feasible $\omega$ such that

$$\omega_{k_1} - \omega_{k_2} = \frac{1}{2}\left(\min_{q\in Q}\{q_{k_2} - q_{k_1}\} + \max_{p\in P}\{p_{k_2} - p_{k_1}\}\right),$$

which satisfies the inequality (51). For this $\omega$, the equalities in (52) hold, and $z$ reaches its the optimal value (50). $\square$

THEOREM 2.6. *Suppose $P$ and $Q$ are two finite sets in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. If for all $p \in P$ and for all $q \in Q$, the indices $i(p)$ and $i(q)$ are respectively constants, say $i_P$ and $i_Q$, and for all $\xi \in P \cup Q$, $j(\xi)$ is a constant, say $j$, then the linear programming (3)–(6) has a feasible solution if and only if*

$$\max_{q\in Q}\{q_{i_P} - q_j\} \ \le\ \min_{p\in P}\{p_{i_P} - p_j\} \tag{53}$$

*and*

$$\max_{p\in P}\{p_{i_Q} - p_j\} \ \le\ \min_{q\in Q}\{q_{i_Q} - q_j\}. \tag{54}$$

*If a feasible solution exists, then the optimal value $z$ is given by*

$$\min\{\ \min_{p\in P}\{p_{i_P} - p_j\} + \min_{q\in Q}\{q_j - q_{i_P}\},\ \min_{q\in Q}\{q_{i_Q} - q_j\} + \min_{p\in P}\{p_j - p_{i_Q}\}\ \}. \tag{55}$$

PROOF. For any $p \in P$, by the assumptions $i(p) = i_P$ and $j(p) = j$, and by (5)–(6), we have:

$$\omega_j - \omega_{i_P} \le p_{i_P} - p_j, \text{ and } \omega_l - \omega_j \le p_j - p_l,\ \ \forall l \ne i_P, j.$$

By (1), $i_P \ne i_Q$. By the definition of $i(q)$ and $j(q)$, $i(q) \ne j(q)$, and hence $i_Q \ne j$. So, there exists $l \ne i_P, j$ such that $l = i_Q$. Similarly, for any $q \in Q$, we have

$$\omega_j - \omega_{i_Q} \le q_{i_Q} - q_j, \text{ and } \omega_l - \omega_j \le q_j - q_l,\ \ \forall l \ne i_Q, j,$$

and there exists $l \neq i_Q, j$ such that $l = i_P$. So we have

$$\forall p \in P,\ \omega_j - \omega_{i_P} \leq p_{i_P} - p_j, \text{ and } \omega_{i_Q} - \omega_j \leq p_j - p_{i_Q}, \tag{56}$$

and

$$\forall q \in Q,\ \omega_j - \omega_{i_Q} \leq q_{i_Q} - q_j, \text{ and } \omega_{i_P} - \omega_j \leq q_j - q_{i_P}. \tag{57}$$

Therefore,

$$\max_{q\in Q}\{q_{i_P} - q_j\} \leq \omega_j - \omega_{i_P} \leq \min_{p\in P}\{p_{i_P} - p_j\} \tag{58}$$

and

$$\max_{p\in P}\{p_{i_Q} - p_j\} \leq \omega_j - \omega_{i_Q} \leq \min_{q\in Q}\{q_{i_Q} - q_j\}. \tag{59}$$

So, if the linear programming (3)–(6) has a feasible solution, then we have (53) and (54).

On the other hand, if we have (53) and (54), then there exist real numbers $\omega_j, \omega_{i_P}$ and $\omega_{i_Q}$ such that (58) and (59) hold, and hence (56) and (57) hold. For $l \neq i_P, j, i_Q$, there always exist sufficiently small numbers for $\omega_l$ such that the inequality (6) holds. So inequalities (53) and (54) guarantee the feasibility of the inequalities (5) and (6). Notice that once (5) and (6) are feasible, there is always a non-negative number $z$ such that (4) holds. So, if we have (53) and (54), then the linear programming (3)–(6) has a feasible solution.

If a feasible solution exists, then by (4) and the first inequality in (58),

$$z \leq \min_{p\in P}\{p_{i_P} - p_j\} + \omega_{i_P} - \omega_j \leq \min_{p\in P}\{p_{i_P} - p_j\} + \min_{q\in Q}\{q_j - q_{i_P}\}, \text{ and}$$

by (4) and the first inequality in (59),

$$z \leq \min_{q\in Q}\{q_{i_Q} - q_j\} + \omega_{i_Q} - \omega_j \leq \min_{q\in Q}\{q_{i_Q} - q_j\} + \min_{p\in P}\{p_j - p_{i_Q}\}.$$

So the maximum $z$ is given by (55), and this optimal value is reached when either $\omega_{i_P} - \omega_j = \min_{q\in Q}\{q_j - q_{i_P}\}$, or $\omega_{i_Q} - \omega_j = \min_{p\in P}\{p_j - p_{i_Q}\}$. □

2.2. *Proofs in Section 4.2 (soft margin tropical SVMs).* Proposition 2.7, Theorem 2.8, Proposition 2.9 and Corollary 2.10 below are respectively Proposition 4.2, Theorem 4.6, Proposition 4.3 and Corollary 4.2 in the main manuscript.

PROPOSITION 2.7. *Given two sets $P$ and $Q$ in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$ ($n = |P| = |Q|$), assume two sets of indices $\mathcal{I}$ and $\mathcal{J}$ satisfy the conditions (1)–(2). Then, the linear programming (7)–(11) has infinitely many feasible solutions.*

PROOF. In fact, for any $\omega^* \in \mathbb{R}^n$, let $z^* = 0$, and for any $\xi \in P \cup Q$, let

$$\alpha^*_\xi = \beta^*_\xi = \max\{0, \left(\omega^*_{j(\xi)} - \omega^*_{i(\xi)}\right) - \left(\xi_{i(\xi)} - \xi_{j(\xi)}\right)\},$$

$$\gamma^*_{\xi,l} = \max\{0, \left(\omega^*_l - \omega^*_{j(\xi)}\right) - \left(\xi_{j(\xi)} - \xi_l\right)\}, \text{ for any } l \neq i(\xi), j(\xi).$$

It is straightforward to check that $(z^*; \alpha^*; \beta^*; \gamma^*; \omega^*) \in \mathbb{R}^{2dn+d+1}$ satisfies the inequalities (8)–(11). □

THEOREM 2.8. *Given two finite sets $P$ and $Q$ in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$ ($n = |P| = |Q|$), assume two sets of indices $\mathcal{I}$ and $\mathcal{J}$ satisfy the conditions (1)–(2). If both $P$ and $Q$ are non-empty, and if $\mathcal{C} \geq 1$, then the objective function in the linear programming (7)–(11) is upper bounded for any feasible solution $(z; \alpha; \beta; \gamma; \omega) \in \mathbb{R}^{2dn+d+1}$, which means the maximum of the objective function is a finite real number.*

PROOF. Since both $P$ and $Q$ are non-empty, pick $p$ and $q$ from $P$ and $Q$ respectively. By the assumptions (1)–(2), we know $i(p) \neq j(p)$, $i(q) \neq j(q)$, and $i(p) \neq i(q)$. Below, we prove the conclusion for the two cases: $i(p) \neq j(q)$ and $i(p) = j(q)$.

**(Case A).** If $i(p) \neq j(q)$, then by (10), we have

$$\omega_{i(p)} - \omega_{j(q)} \leq q_{j(q)} - q_{i(p)} + \gamma_{q,i(p)}. \tag{60}$$

**(Case A.1).** If $j(q) = j(p)$, then (60) becomes

$$\omega_{i(p)} - \omega_{j(p)} \leq q_{j(p)} - q_{i(p)} + \gamma_{q,i(p)}. \tag{61}$$

Then, by (8) and (61),

$$z \leq p_{i(p)} - p_{j(p)} + q_{j(p)} - q_{i(p)} + \gamma_{q,i(p)} + \alpha_p.$$

So, if $\mathcal{C} \geq 1$, then we have an upper bound for the objective function

$$z - \mathcal{C} \sum_{\xi \in P \cup Q} \left( \alpha_\xi + \beta_\xi + \sum_{l \neq i(\xi), j(\xi)} \gamma_{\xi,l} \right) \leq z - \alpha_p - \gamma_{q,i(p)} \leq p_{i(p)} - p_{j(p)} + q_{j(p)} - q_{i(p)}.$$

**(Case A.2).** If $j(q) \neq j(p)$, then by (10), we have

$$\omega_{j(q)} - \omega_{j(p)} \leq p_{j(p)} - p_{j(q)} + \gamma_{p,j(q)}. \tag{62}$$

By summing up (60) and (62), we have

$$\omega_{i(p)} - \omega_{j(p)} \leq q_{j(q)} - q_{i(p)} + \gamma_{q,i(p)} + p_{j(p)} - p_{j(q)} + \gamma_{p,j(q)}. \tag{63}$$

So if $\mathcal{C} \geq 1$, then by (8) and (63), we have an upper bound for the objective function

$$z - \mathcal{C} \sum_{\xi \in P \cup Q} \left( \alpha_\xi + \beta_\xi + \sum_{l \neq i(\xi), j(\xi)} \gamma_{\xi,l} \right) \leq q_{j(q)} - q_{i(p)} + p_{i(p)} - p_{j(q)}.$$

**(Case B).** If $i(p) = j(q)$, then by (8),

$$\omega_{i(p)} - \omega_{i(q)} \leq q_{i(q)} - q_{i(p)} + \alpha_q. \tag{64}$$

**(Case B.1).** If $i(q) = j(p)$, then (64) becomes

$$\omega_{i(p)} - \omega_{j(p)} \leq q_{j(p)} - q_{i(p)} + \alpha_q. \tag{65}$$

So if $\mathcal{C} \geq 1$, then by (8) and (65), we have an upper bound for the objective function

$$z - \mathcal{C} \sum_{\xi \in P \cup Q} \left( \alpha_\xi + \beta_\xi + \sum_{l \neq i(\xi), j(\xi)} \gamma_{\xi,l} \right) \leq z - \alpha_q - \alpha_p \leq q_{j(p)} - q_{i(p)} + p_{i(p)} - p_{j(p)}.$$

**(Case B.2).** If $i(q) \neq j(p)$,then by (10), we have

$$\omega_{i(q)} - \omega_{j(p)} \leq p_{j(p)} - p_{i(q)} + \gamma_{p,i(q)}. \tag{66}$$

By summing up (64) and (66),

$$\omega_{i(p)} - \omega_{j(p)} \leq q_{i(q)} - q_{i(p)} + \alpha_q + p_{j(p)} - p_{i(q)} + \gamma_{p,i(q)}. \tag{67}$$

So if $\mathcal{C} \geq 1$, then by (8) and (67), we have an upper bound for the objective function

$$z - \mathcal{C} \sum_{\xi \in P \cup Q} \left( \alpha_\xi + \beta_\xi + \sum_{l \neq i(\xi), j(\xi)} \gamma_{\xi,l} \right) \leq z - \alpha_q - \alpha_p - \gamma_{p,i(q)} \leq q_{i(q)} - q_{i(p)} + p_{i(p)} - p_{i(q)}.$$

$\square$

PROPOSITION 2.9. *For all $p \in P$, assume $i(p)$ and $j(p)$ are constants, say $i_P$ and $j_P$. For all $q \in Q$, assume $i(q)$ and $j(q)$ are constants, say $i_Q$ and $j_Q$. If $(z^*;\alpha^*;\beta^*;\gamma^*;\omega^*)$ is an optimal solution to the linear programming (7)–(11), then for any $l \neq i_P, j_P, i_Q, j_Q$, we have $\gamma^*_{\xi,l} = 0$ for all $\xi \in P \cup Q$.*

PROOF. Assume that there exists $\tilde{l} \neq i_P, j_P, i_Q, j_Q$ such that $\gamma^*_{\xi,\tilde{l}} > 0$ for some $\xi \in P \cup Q$. By (10),

$$\omega^*_{\tilde{l}} - \omega^*_{j(\xi)} \;\leq\; \xi_{j(\xi)} - \xi_{\tilde{l}} + \gamma^*_{\xi,\tilde{l}}.$$

Let $\hat{\omega}^*_{\tilde{l}} = \omega^*_{\tilde{l}} - \gamma^*_{\xi,\tilde{l}}$, and let $\hat{\gamma}^*_{\xi,\tilde{l}} = 0$. We replace the coordinates $\omega^*_{\tilde{l}}$ and $\gamma^*_{\xi,\tilde{l}}$ of the vector $(z^*;\alpha^*;\beta^*;\gamma^*;\omega^*)$ with $\hat{\omega}^*_{\tilde{l}}$ and $\hat{\gamma}^*_{\xi,\tilde{l}}$, and we call the resulting vector $(z^*;\alpha^*;\beta^*;\hat{\gamma}^*;\hat{\omega}^*)$. Then $(z^*;\alpha^*;\beta^*;\hat{\gamma}^*;\hat{\omega}^*)$ is still a feasible solution, for which the objective function has the value

$$z^* - \mathcal{C} \sum_{\xi \in P \cup Q} \left( \alpha^*_\xi + \beta^*_\xi + \sum_{l \neq i(\xi), j(\xi)} \gamma^*_{\xi,l} \right) + \gamma^*_{\xi,\tilde{l}}.$$

That means $(z^*;\alpha^*;\beta^*;\hat{\gamma}^*;\hat{\omega}^*)$ gives a larger function value to the objective function, which is a contradiction to the fact that the objective function reaches its maximum at $(z^*;\alpha^*;\beta^*;\gamma^*;\omega^*)$. $\square$

COROLLARY 2.10. *For all $p \in P$, assume $i(p)$ and $j(p)$ are constants, say $i_P$ and $j_P$. For all $q \in Q$, assume $i(q)$ and $j(q)$ are constants, say $i_Q$ and $j_Q$. Then any point in $P$ is located in the closed sector $\overline{S}^{i_P}_{\omega^*}$ or in $\overline{S}^{j_P}_{\omega^*}$, and any point in $Q$ is located in the closed sector $\overline{S}^{i_Q}_{\omega^*}$ or in $\overline{S}^{j_Q}_{\omega^*}$ of an optimal tropical hyperplane $H_{\omega^*}$.*

PROOF. Suppose $(z^*;\alpha^*;\beta^*;\gamma^*;\omega^*)$ is a feasible solution to the linear programming (7)–(11) such that the objective function reaches its maximum. By Proposition 2.9, for any $p \in P$, we have

$$\omega^*_l - \omega^*_{j_P} \;\leq\; p_{j_P} - p_l \;\Leftrightarrow\; (p + \omega^*)_l \;\leq\; (p + \omega^*)_{j_P}$$

for any $l \neq i_P, j_P$. Then the maximum coordinate of vector $p + \omega^*$ can be indexed by $i_P$ or $j_P$. So, any point in $P$ is located in the closed sector $\overline{S}^{i_P}_{\omega^*}$ or in $\overline{S}^{j_P}_{\omega^*}$. Similarly, we can show that any point in $Q$ is located in the closed sector $\overline{S}^{i_Q}_{\omega^*}$ or in $\overline{S}^{j_Q}_{\omega^*}$. $\square$

**3. Algorithms.** For the four cases (**Case 1**)–(**Case 4**) list in Section 4.2 (see the main manuscript), we respectively simplify (7)–(11) as four linear programing problems (**LP 1**)-(**LP 4**) (for (**Case 2**), we only show the simplified linear programming problem for $i_P = j_Q$ and $i_Q \neq j_P$).

$$
\begin{aligned}
\max_{(z;\alpha;\beta;\gamma)\in\mathbb{R}^{8n+1}_{\geq 0}} \; & z - \sum_{\xi\in P\cup Q} \left(\alpha_\xi + \beta_\xi\right) \\
& - \sum_{p\in P,\, l=i_Q,j_Q} \gamma_{p,l} - \sum_{q\in Q,\, l=i_P,j_P} \gamma_{q,l} \\
\text{s.t. } \forall p\in P, \; & z + p_{j_P} + \omega_{j_P} - p_{i_P} - \omega_{i_P} \leq \alpha_p, \\
& \omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P} + \beta_p, \\
& \omega_{i_Q} - \omega_{j_P} \leq p_{j_P} - p_{i_Q} + \gamma_{p,i_Q}, \\
& \omega_{j_Q} - \omega_{j_P} \leq p_{j_P} - p_{j_Q} + \gamma_{p,j_Q}, \\
& \omega_l - \omega_{j_P} \leq p_{j_P} - p_l \\
& \qquad (\forall l \neq i_P, j_P, i_Q, j_Q), \\
\forall q\in Q, \; & z + q_{j_Q} + \omega_{j_Q} - q_{i_Q} - \omega_{i_Q} \leq \alpha_q, \\
& \omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q} + \beta_q, \\
& \omega_{i_P} - \omega_{j_Q} \leq q_{j_Q} - q_{i_P} + \gamma_{q,i_P}, \\
& \omega_{j_P} - \omega_{j_Q} \leq q_{j_Q} - q_{j_P} + \gamma_{q,j_P}, \\
& \omega_l - \omega_{j_Q} \leq q_{j_Q} - q_l \\
& \qquad (\forall l \neq i_P, j_P, i_Q, j_Q).
\end{aligned}
\tag{LP 1}
$$

$$
\begin{aligned}
\max_{(z;\alpha;\beta;\gamma)\in\mathbb{R}^{6n+1}_{\geq 0}} \; & z - \sum_{\xi\in P\cup Q} \left(\alpha_\xi + \beta_\xi\right) \\
& - \sum_{p\in P} \gamma_{p,i_Q} - \sum_{q\in Q} \gamma_{q,j_P} \\
\text{s.t. } \forall p\in P, \; & z + p_{j_P} + \omega_{j_P} - p_{i_P} - \omega_{i_P} \leq \alpha_p, \\
& \omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P} + \beta_p, \\
& \omega_{i_Q} - \omega_{j_P} \leq p_{j_P} - p_{i_Q} + \gamma_{p,i_Q}, \\
& \omega_l - \omega_{j_P} \leq p_{j_P} - p_l \\
& \qquad (\forall l \neq i_P, j_P, i_Q), \\
\forall q\in Q, \; & z + q_{j_Q} + \omega_{j_Q} - q_{i_Q} - \omega_{i_Q} \leq \alpha_q, \\
& \omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q} + \beta_q, \\
& \omega_{j_P} - \omega_{j_Q} \leq q_{j_Q} - q_{j_P} + \gamma_{q,j_P}, \\
& \omega_l - \omega_{j_Q} \leq q_{j_Q} - q_l \\
& \qquad (\forall l \neq i_P, j_P, i_Q).
\end{aligned}
\tag{LP 2}
$$

$$
\begin{aligned}
\max_{(z;\alpha;\beta)\in\mathbb{R}^{4n+1}_{\geq 0}} \; & z - \sum_{\xi\in P\cup Q} \left(\alpha_\xi + \beta_\xi\right) \\
\text{s.t. } \forall p\in P, \; & z + p_{k_2} + \omega_{k_2} - p_{k_1} - \omega_{k_1} \leq \alpha_p, \\
& \omega_{k_2} - \omega_{k_1} \leq p_{k_1} - p_{k_2} + \beta_p, \\
& \omega_l - \omega_{k_2} \leq p_{k_2} - p_l \\
& \qquad (\forall l \neq k_1, k_2), \\
\forall q\in Q, \; & z + q_{k_1} + \omega_{k_1} - q_{k_2} - \omega_{k_2} \leq \alpha_q, \\
& \omega_{k_1} - \omega_{k_2} \leq q_{k_2} - q_{k_1} + \beta_q, \\
& \omega_l - \omega_{k_1} \leq q_{k_1} - q_l \\
& \qquad (\forall l \neq k_1, k_2).
\end{aligned}
\tag{LP 3}
$$

$$
\begin{aligned}
\max_{(z;\alpha;\beta;\gamma)\in\mathbb{R}^{6n+1}_{\geq 0}} \; & z - \sum_{\xi\in P\cup Q} \left(\alpha_\xi + \beta_\xi\right) \\
& - \sum_{p\in P} \gamma_{p,i_Q} - \sum_{q\in Q} \gamma_{q,i_P} \\
\text{s.t. } \forall p\in P, \; & z + p_{j_P} + \omega_{j_P} - p_{i_P} - \omega_{i_P} \leq \alpha_p, \\
& \omega_{j_P} - \omega_{i_P} \leq p_{i_P} - p_{j_P} + \beta_p, \\
& \omega_{i_Q} - \omega_j \leq p_j - p_{i_Q} + \gamma_{p,i_Q}, \\
& \omega_l - \omega_j \leq p_j - p_l \\
& \qquad (\forall l \neq i_P, j, i_Q), \\
\forall q\in Q, \; & z + q_{j_Q} + \omega_{j_Q} - q_{i_Q} - \omega_{i_Q} \leq \alpha_q, \\
& \omega_{j_Q} - \omega_{i_Q} \leq q_{i_Q} - q_{j_Q} + \beta_q, \\
& \omega_{i_P} - \omega_j \leq q_j - q_{i_P} + \gamma_{q,i_P}, \\
& \omega_l - \omega_j \leq q_j - q_l \\
& \qquad (\forall l \neq i_P, j, i_Q).
\end{aligned}
\tag{LP 4}
$$

We develop Algorithms 1–4 according to the simplified linear programmings (**LP 1**)–(**LP 4**) respectively

The input of each algorithm includes two sets $P$ and $Q$ ($P \cap Q = \emptyset$ and $|P| = |Q| = n$), a test set $T$ and indices $i_P, j_P, i_Q, j_Q$ for formulating the corresponding linear programming. The sets $P$ and $Q$ are associated with a training dataset $\{(x^{(1)}, y_1), \ldots, (x^{(2n)}, y_{2n})\} \subset \mathbb{R}^d/\mathbb{R}\mathbf{1} \times \{0,1\}$. Here, in these algorithms, we simply call $P$ and $Q$ training sets. The test set $T$ is a finite subset of $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. The indices $i_P, j_P, i_Q, j_Q$ are all from $\{1, \ldots, d\}$ and they satisfy $i_P \neq j_P$, $i_Q \neq j_Q$, and $i_P \neq i_Q$. There are two main steps in each algorihm:

**Step 1.** In Algorithms 1–4, for the input sets $P$ and $Q$ and indices $i_P, j_P, i_Q, j_Q$, we solve the corresponding linear programming ((**LP 1**)–(**LP 4**) respectively) and obtain the normal

**Algorithm 1: Tropical Classifier via LP 1**

**input** : Training sets: $P$, $Q$; Test set: $T$; Indices: pairwise distinct $i_P, i_Q, j_P, j_Q$; Threshold: $\eta > 0$; Parameter: $C > 0$ (for the multispecies coalescent process)

**output**: Optimal normal vector: $\omega$; A partition of $T$: $\tilde{P}$, $\tilde{Q}$ such that $\tilde{P} \cup \tilde{Q} = T$ and $\tilde{P} \cap \tilde{Q} = \emptyset$

1 Solve the linear programming (**LP 1**) for input data sets $P, Q$ and indices $i_P, i_Q, j_P, j_Q$

2 $\omega \leftarrow$ optimal $\omega$ such that the objective function in (**LP 1**) reaches its optimal value

3 $\tilde{P} \leftarrow \emptyset, \quad \tilde{Q} \leftarrow \emptyset$

4 **for** *each* $t \in T$ **do**

5     $\mathcal{I} \leftarrow$ the set of indices $\{i | \omega_i + t_i = \max\{\omega_k + t_k | 1 \leq k \leq d\}, 1 \leq i \leq d\}$

6     $\mathcal{I}^* \leftarrow \mathcal{I} \cap \{i_P, i_Q, j_P, j_Q\}$

7     **if** $C \leq \eta$ **then**

8         **if** $\mathcal{I}^* = \{i_P\}, \{j_P\}, \{i_P, j_P\}, \{i_P, i_Q\}, \{j_P, i_Q\}, \{j_P, j_Q\}, \{j_P, i_Q, j_Q\}$, or $\{i_P, j_P, i_Q, j_Q\}$ **then** Add $t$ into $\tilde{P}$

9         **if** $\mathcal{I}^* = \{i_Q\}, \{j_Q\}, \{i_Q, j_Q\}, \{i_P, j_Q\}, \{i_P, j_P, i_Q\}, \{i_P, j_P, j_Q\}, \{i_P, i_Q, j_Q\}$, or $\emptyset$ **then** Add $t$ into $\tilde{Q}$

10     **if** $C > \eta$ **then**

11         **if** $\mathcal{I}^* = \{i_P\}, \{j_P\}, \{i_P, j_P\}, \{i_P, i_Q\}, \{j_P, j_Q\}, \{i_P, i_Q, j_Q\}, \{j_P, i_Q, j_Q\}$, or $\{i_P, j_P, i_Q, j_Q\}$ **then** Add $t$ into $\tilde{P}$

12         **if** $\mathcal{I}^* = \{i_Q\}, \{j_Q\}, \{i_Q, j_Q\}, \{i_P, j_Q\}, \{j_P, i_Q\}, \{i_P, j_P, i_Q\}, \{i_P, j_P, j_Q\}$, or $\emptyset$ **then** Add $t$ into $\tilde{Q}$

13 **return** $\omega$, $\tilde{P}$, $\tilde{Q}$

vector $\omega$ of an optimal tropical hyperplane, which separates the two categories of data $P$ and $Q$.

**Step 2.** After that, for each point $t$ from the test set $T$, we add $t$ into the set $\tilde{P}$ or $\tilde{Q}$ (that means we classify the point $t$ as the category $P$ or $Q$) according to which sector of $H_\omega$ the point $t$ is located in. As a result, the test set $T$ will be divided into two subsets $\tilde{P}$ and $\tilde{Q}$, and the output of each algorithm is the optimal normal vector $\omega$ and a partition of the test set: $\tilde{P}$ and $\tilde{Q}$.

Below, we give more details for the above Step 2. The key of this step is to decide which category a point $t$ from the test set should go once we have the optimal $H_\omega$ solved from the linear programming. The point $t$ might be located in a closed sector, or on an intersection of many different closed sectors of $H_\omega$ (in comparison, for a soft margin classical SVM, a point from the test set might be simply in one of two open half-spaces determined by an optimal hyperplane, or on the hyperplane). Remark that for a tropical hyperplane in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$, there are $d$ closed sectors and $2^d - d - 1$ possible intersections of different closed sectors. So far, we do not prove any criteria on how to classify a point according to its location. Here, according to substantial experiments on simulated data generated by the multispecies coalescent process, we propose an effective strategy in Algorithms 1–4 as follows. Since our input training data $P$ and $Q$ are generated by the multispecies coalescent process, we also input two numbers $C$ and $\eta$ to these algorithms, where $C$ denotes the ratio of the depth of the species tree to the effective population size in the multispecies coalescent process, and $\eta$ is a threshold (experiments show that a real number between 4 and 5 is a good choice for $\eta$). In each algorithm, for the input data $P$ and $Q$, we provide two ways to classify a point from the test set according to the relative values of $C$ and $\eta$. For instance, in Algorithm 3, the variable $\mathcal{I}^*$ in Line 6 has 4 possible values: $\{k_1, k_2\}$, $\{k_1\}$, $\{k_2\}$, or $\emptyset$. That respectively means the current point $t \in T$ read by Line 4 is located on the intersection of $\overline{S}_\omega^{k_1}$ and $\overline{S}_\omega^{k_2}$, in the difference $\overline{S}_\omega^{k_1} \backslash \overline{S}_\omega^{k_2}$, in the difference $\overline{S}_\omega^{k_2} \backslash \overline{S}_\omega^{k_1}$, or other cases. When the input $C$ is not larger than the input $\eta$, we apply one method to classify $t$ (see Lines 8–9), and when $C$ is larger than $\eta$, we apply another method (see Lines 11–12). The other three algorithms are similarly designed. In our experiments (Section 5 in the main manuscript), we set $\eta = 4.8$.

**Algorithm 2: Tropical Classifier via LP 2**

**input** : Training sets: $P$, $Q$; Test set: $T$; Indices: pairwise distinct $i_P$, $i_Q$, $j_P$ ($j_Q = i_P$); Threshold: $\eta > 0$; Parameter: $C > 0$ (for the multispecies coalescent process)
**output**: Optimal normal vector: $\omega$; A partition of $T$: $\tilde{P}$, $\tilde{Q}$ such that $\tilde{P} \cup \tilde{Q} = T$ and $\tilde{P} \cap \tilde{Q} = \emptyset$

Solve the linear programming (LP 2) for input data sets $P$, $Q$ and indices $i_P$, $i_Q$, $j_P$
$\omega \leftarrow$ optimal $\omega$ such that the objective function in (LP 2) reaches its optimal value
$\tilde{P} \leftarrow \emptyset$, $\tilde{Q} \leftarrow \emptyset$
**for** *each* $t \in T$ **do**
  $\mathcal{I} \leftarrow$ the set of indices $\{i | \omega_i + t_i = \max\{\omega_k + t_k | 1 \leq k \leq d\}, 1 \leq i \leq d\}$
  $\mathcal{I}^* \leftarrow \mathcal{I} \cap \{i_P, i_Q, j_P\}$
  **if** $C \leq \eta$ **then**
    **if** $\mathcal{I}^* = \{i_P\}, \{j_P\}, \{i_Q\}$, or $\{i_Q, j_P\}$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{i_P, j_P\}, \{i_P, i_Q\}, \{i_P, i_Q, j_P\}$, or $\emptyset$ **then** Add $t$ into $\tilde{Q}$
  **if** $C > \eta$ **then**
    **if** $\mathcal{I}^* = \{i_P\}, \{i_P, j_P\}, \{i_P, i_Q, j_P\}$, or $\emptyset$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{j_P\}, \{i_Q\}, \{i_P, i_Q\}$, or $\{i_Q, j_P\}$ **then** Add $t$ into $\tilde{Q}$
**return** $\omega$, $\tilde{P}$, $\tilde{Q}$

**Algorithm 3: Tropical Classifier via LP 3**

**input** : Training sets: $P$, $Q$; Test set: $T$; Indices: distinct $k_1$, $k_2$ ($i_P = j_Q = k_1$, $i_Q = j_P = k_2$); Threshold: $\eta > 0$; Parameter: $C > 0$ (for the multispecies coalescent process)
**output**: Optimal normal vector: $\omega$; A partition of $T$: $\tilde{P}$, $\tilde{Q}$ such that $\tilde{P} \cup \tilde{Q} = T$ and $\tilde{P} \cap \tilde{Q} = \emptyset$

Solve the linear programming (LP 3) for input data sets $P$, $Q$ and indices $k_1$, $k_2$
$\omega \leftarrow$ optimal $\omega$ such that the objective function in (LP 3) reaches its optimal value
$\tilde{P} \leftarrow \emptyset$, $\tilde{Q} \leftarrow \emptyset$
**for** *each* $t \in T$ **do**
  $\mathcal{I} \leftarrow$ the set of indices $\{i | \omega_i + t_i = \max\{\omega_k + t_k | 1 \leq k \leq d\}, 1 \leq i \leq d\}$
  $\mathcal{I}^* \leftarrow \mathcal{I} \cap \{k_1, k_2\}$
  **if** $C \leq \eta$ **then**
    **if** $\mathcal{I}^* = \{k_1\}$, or $\{k_1, k_2\}$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{k_2\}$, or $\emptyset$ **then** Add $t$ into $\tilde{Q}$
  **if** $C > \eta$ **then**
    **if** $\mathcal{I}^* = \{k_1\}$, or $\emptyset$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{k_2\}$, or $\{k_1, k_2\}$ **then** Add $t$ into $\tilde{Q}$
**return** $\omega$, $\tilde{P}$, $\tilde{Q}$

**Algorithm 4: Tropical Classifier via LP 4**

**input** : Training sets: $P$, $Q$; Test set: $T$; Indices: pairwise distinct $i_P$, $i_Q$, $j$ ($j_P = j_Q = j$); Threshold: $\eta > 0$; Parameter: $C > 0$ (for the multispecies coalescent process)
**output**: Optimal normal vector: $\omega$; A partition of $T$: $\tilde{P}$, $\tilde{Q}$ such that $\tilde{P} \cup \tilde{Q} = T$ and $\tilde{P} \cap \tilde{Q} = \emptyset$

Solve the linear programming (LP 4) for input data sets $P$, $Q$ and indices $i_P$, $i_Q$, $j$
$\omega \leftarrow$ optimal $\omega$ such that the objective function in (LP 4) reaches its optimal value
$\tilde{P} \leftarrow \emptyset$, $\tilde{Q} \leftarrow \emptyset$
**for** *each* $t \in T$ **do**
  $\mathcal{I} \leftarrow$ the set of indices $\{i | \omega_i + t_i = \max\{\omega_k + t_k | 1 \leq k \leq d\}, 1 \leq i \leq d\}$
  $\mathcal{I}^* \leftarrow \mathcal{I} \cap \{i_P, i_Q, j\}$
  **if** $C \leq \eta$ **then**
    **if** $\mathcal{I}^* = \{i_Q\}, \{i_P, i_Q\}, \{i_P, j\}$, or $\{j\}$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{i_P\}, \{i_Q, j\}, \{i_P, i_Q, j\}$, or $\emptyset$ **then** Add $t$ into $\tilde{Q}$
  **if** $C > \eta$ **then**
    **if** $\mathcal{I}^* = \{i_P\}, \{i_P, i_Q\}, \{i_P, i_Q, j\}$, or $\emptyset$ **then** Add $t$ into $\tilde{P}$
    **if** $\mathcal{I}^* = \{i_Q\}, \{i_P, j\}, \{i_Q, j\}$, or $\{j\}$ **then** Add $t$ into $\tilde{Q}$
**return** $\omega$, $\tilde{P}$, $\tilde{Q}$