Efficient Point Process Inference for Large-Scale Object Detection
Supplemental Material

Trung T. Pham, Seyed Hamid Rezatofighi, Ian Reid and Tat-Jun Chin
School of Computer Science
The University of Adelaide
{trung.pham,hamid.rezatofighi,ian.reid,tjchin}@adelaide.edu.au

Proof of Proposition 1

For convenience, we restate the energy functions and the proposition in the main paper. Consider the following two energy functions:

$$E(X) = \sum_{i=1}^{N} D(u_i)x_i + \sum_{x_i \sim x_j} V(u_i, u_j)x_i x_j. \tag{1}$$

$$V(u_i, u_j) = \begin{cases} g(\mathcal{R}(u_i, u_j)) & \text{if } \mathcal{R}(u_i, u_j) < T_o \\ K & \text{if } \mathcal{R}(u_i, u_j) \geq T_o, \end{cases} \tag{2}$$

where $g(\mathcal{R}(u_i, u_j)) \geq 0$, and K is a large constant number, i.e., $K \gg g(\mathcal{R}(u_i, u_j))$ and $K \gg \sup(D(u_i))\,\forall u_i, u_j$.

$$\hat{E}(X) = \sum_{i=1}^{N} D(u_i)x_i + \sum_{x_i \sim x_j} \hat{V}(u_i, u_j)x_i x_j. \tag{3}$$

$$\hat{V}(u_i, u_j) = \begin{cases} g(\mathcal{R}(u_i, u_j)) & \text{if } \mathcal{R}(u_i, u_j) < T_o \\ \alpha & \text{if } \mathcal{R}(u_i, u_j) \geq T_o, \end{cases} \tag{4}$$

where $\alpha = \max(|D(u_i)|, |D(u_j)|) + \epsilon$; ϵ is a small positive number. For notational simplicity, we denote $D(x_i) = D(u_i)x_i$, $V(x_i, x_j) = V(u_i, u_j)x_i x_j$, $\hat{V}(x_i, x_j) = \hat{V}(u_i, u_j)x_i x_j$, and $\mathcal{R}(x_i, x_j) = \mathcal{R}(u_i, u_j)x_i x_j$.

Lemma 1. *If X^* is the global optimum of $E(X)$, $\mathcal{R}(x_i, x_j) < T_o, \forall\, x_i \sim x_j, x_i, x_j \in X^*$.*

Proof. If there exists a pair $x_i \sim x_j, x_i, x_j \in X^*$ such that $\mathcal{R}(x_i, x_j) \geq T_o$, then

$$E(X^*) > 0 = E(\mathbf{0}) \tag{5}$$

where $\mathbf{0}$ is a all-zero vector of length N. This is because $V(x_i, x_j) = K$ ($K \gg g(\mathcal{R}(u_i, u_j))$ and $K \gg \sup(D(u_i))\,\forall u_i, u_j$). This contradicts the fact that X^* is a global minimizer of $E(X)$. □

Lemma 2. *If X^* is the global optimum of $\hat{E}(X)$, $\mathcal{R}(x_i, x_j) < T_o, \forall\, x_i \sim x_j, x_i, x_j \in X^*$.*

Proof. If there exists a pair $x_i \sim x_j, x_i, x_j \in X^*$ such that $\mathcal{R}(x_i, x_j) \geq T_o$, then construct a solution $Z = X^*$, except $z_i = 0$ (*i.e.* $\forall j \neq i,\ z_j = x_j, z_i = 0$). Note that $x_i = x_j = 1$ since $\mathcal{R}(x_i, x_j) \geq T_o$. Consider

$$\begin{aligned}\hat{E}(X^*) - \hat{E}(Z) &= D(x_j) + \hat{V}(x_i, x_j) + \sum_{\substack{x_k \sim x_i \\ k \neq j}} \hat{V}(x_i, x_k) \\ &= D(x_j) + \max(|D(x_i)|, |D(x_j)|) + \epsilon + \sum_{\substack{x_k \sim x_i \\ k \neq j}} \hat{V}(x_i, x_k) > 0. \end{aligned} \tag{6}$$

Therefore $\hat{E}(X^*) > \hat{E}(Z)$, which contradicts X^* is a global minimizer of $\hat{E}(X)$.

If X^* contains a set of pairs $\mathcal{E} = \{x_i \sim x_j \mid i > j,\ x_i, x_j \in X^*,\ \mathcal{R}(x_i, x_j) \geq T_o\}$, let define a solution $Z = X^*$, except that $z_i = 0, \forall\, x_i \sim x_j \in \mathcal{E}$, then consider

$$\begin{aligned}\hat{E}(X^*) - \hat{E}(Z) &= \sum_{\substack{x_j \\ x_i \sim x_j \in \mathcal{E}}} D(x_j) + \sum_{x_i \sim x_j \in \mathcal{E}} \hat{V}(x_i, x_j) + \sum_{\substack{x_k \sim x_i \\ x_k \sim x_i \notin \mathcal{E}}} \hat{V}(x_k, x_i) \\ &= \sum_{\substack{x_j \\ x_i \sim x_j \in \mathcal{E}}} D(x_j) + \sum_{x_i \sim x_j \in \mathcal{E}} (\max(|D(x_i)|, |D(x_j)|) + \epsilon) + \sum_{\substack{x_k \sim x_i \\ x_k \sim x_i \notin \mathcal{E}}} \hat{V}(x_k, x_i) > 0.\end{aligned} \tag{7}$$

As a result $\hat{E}(X^*) > \hat{E}(Z)$, which contradicts X^* is a global minimizer of $\hat{E}(X)$. □

Lemma 3. *For all X, if $\forall\, x_i \sim x_j, x_i, x_j \in X,\ \mathcal{R}(x_i, x_j) < T_o$, $\hat{E}(X) = E(X)$.*

Proof. As $\mathcal{R}(x_i, x_j) < T_o \Rightarrow \hat{V}(x_i, x_j) = V(x_i, x_j)$. Therefore $\hat{E}(X) = E(X)$. □

Lemma 4. *If X^* is the global optimum of $\hat{E}(X)$, $\hat{E}(X^*) = E(X^*)$.*

Proof. Since X^* is the global optimum of $\hat{E}(X)$, using Lemma 2, we get $\forall\, x_i \sim x_j, x_i, x_j \in X^*,\ \mathcal{R}(x_i, x_j) < T_o$. As a result, $E(X^*) = \hat{E}(X^*)$. □

Proposition 1. *If X^* is the globally minimal solution of the energy function $\hat{E}(X)$, X^* is also the global minimizer of the function $E(X)$, and vice versa.*

Proof.

Typically the proposition 1 says that the two energy functions $E(X)$ and $\hat{E}(X)$ admit the same global minimizer X^*. In fact, assume that X_o^* and X_n^* are the global optimum of the energies $E(X)$ and $\hat{E}(X)$ respectively. Denote $\mathcal{X} = \{X \mid \forall\, x_i \sim x_j, x_i, x_j \in X, \mathcal{R}(x_i, x_j) < T_o\}$. Lemmas 1 and 2 say that $X_o^* \in \mathcal{X}$ and $X_n^* \in \mathcal{X}$. Moreover lemma 3 reveals that $E(X) = \hat{E}(X), \forall X \in \mathcal{X}$. Therefore we can conclude that $X_o^* = X_n^* = X^*$. □

Proof. (By contradiction)

We start proving the first part of the proposition 1. Suppose that X^* is the global optimum of $\hat{E}(X)$, i.e.,

$$\hat{E}(X^*) \leq \hat{E}(X),\ \forall\, X, \tag{8}$$

we want to prove that X^* is also a global minimum of $E(X)$ Assume that X^* is not the global optimum of $E(X)$, then there exists a solution Y such that

$$E(Y) < E(X^*). \tag{9}$$

Now we consider the following scenarios:

1. $\forall\, y_i \sim y_j, y_i, y_j \in Y, \mathcal{R}(y_i, y_j) < T_o$.

   Using Lemma 3, we obtain

$$E(Y) = \hat{E}(Y). \tag{10}$$

   Since $E(Y) \leq E(X^*)$ (see (9)), $\hat{E}(Y) \leq E(Y)$ and $\hat{E}(X^*) = E(X^*)$ (using Lemma 4), we arrive at $\hat{E}(Y) < \hat{E}(X^*)$ which contradicts (8).

2. $\exists\, y_i \sim y_j, y_i, y_j \in Y$ such that $\mathcal{R}(y_i, y_j) \geq T_o$.

   Let define a solution $Z = Y$, except $z_i = 0$ (*i.e.* $\forall j \neq i,\; z_j = y_j, z_i = 0$). Note that $y_i = y_j = 1$ since $\mathcal{R}(y_i, y_j) \geq T_o$. Consider

$$E(Y) - E(Z) = D(y_j) + V(y_i, y_j) + \sum_{\substack{y_k \sim y_i \\ k \neq j}} V(y_i, y_k) = D(y_j) + K + \sum_{\substack{y_k \sim y_i \\ k \neq j}} V(y_i, y_k) > 0. \tag{11}$$

   Therefore $E(Z) < E(Y)$, and $\hat{E}(Z) \leq E(Z) < E(Y) < E(X^*)$. As $\hat{E}(X^*) = E(X^*)$ (using Lemma 4), $\hat{E}(Z) < \hat{E}(X^*)$, which again contradicts (8).

3. Y contains a set of pairs $\mathcal{E} = \{y_i \sim y_j \mid i > j,\; y_i, y_j \in Y,\; \mathcal{R}(y_i, y_j) \geq T_o\}$.

   Similar to the above reasoning, let define a solution $Z = Y$, except that $z_i = 0, \forall\, y_i \sim y_j \in \mathcal{E}$, then consider

$$\begin{aligned} E(Y) - E(Z) &= \sum_{\substack{y_j \\ y_i \sim y_j \in \mathcal{E}}} D(y_j) + \sum_{y_i \sim y_j \in \mathcal{E}} V(y_i, y_j) + \sum_{\substack{y_k \sim y_i \\ y_k \sim y_i \notin \mathcal{E}}} V(y_k, y_i) \\ &= \sum_{\substack{y_j \\ y_i \sim y_j \in \mathcal{E}}} D(y_j) + |\mathcal{E}|K + \sum_{\substack{y_k \sim y_i \\ y_k \sim y_i \notin \mathcal{E}}} V(y_k, y_i) > 0, \end{aligned} \tag{12}$$

   where $|\mathcal{E}|$ is the number of pairs in $\mathcal{E}$. As a result $E(Z) < E(Y)$, and $\hat{E}(Z) \leq E(Z) < E(Y) < E(X^*)$. As $\hat{E}(X^*) = E(X^*)$ (using Lemma 4), $\hat{E}(Z) < \hat{E}(X^*)$, which again contradicts (8).

It can be seen that the above assumption is always wrong, thus X^* is the global optimum of $E(X)$. The second part of the proposition 1 can be proved similarly. Suppose that X^* is the global optimum of $E(X)$, i.e.,

$$E(X^*) \leq E(X),\; \forall\, X, \tag{13}$$

we want to prove that X^* is also a global minimum of $\hat{E}(X)$. Assume that X^* is not the global optimum of $\hat{E}(X)$, then there exists a solution Y such that

$$\hat{E}(Y) < \hat{E}(X^*).$$

If Y contains no pairs $y_i \sim y_j$ such that $\mathcal{R}(y_i, y_j) < T_o$, Lemma 3 says

$$E(Y) = \hat{E}(Y). \tag{14}$$

Since $\hat{E}(Y) < \hat{E}(X^*)$ and $\hat{E}(X^*) \leq E(X^*)$, we arrive at $E(Y) < E(X^*)$ which contradicts (13).

If the solution Y contains a set of pairs $\mathcal{E} = \{y_i \sim y_j \mid i > j,\ y_i, y_j \in Y,\ \mathcal{R}(y_i, y_j) \geq T_o\}$, let define a solution $Z = Y$, except that $z_i = 0, \forall\, y_i \sim y_j \in \mathcal{E}$, then consider

$$\begin{aligned}
\hat{E}(Y) - \hat{E}(Z) &= \sum_{\substack{y_j \\ y_i \sim y_j \in \mathcal{E}}} D(y_j) + \sum_{y_i \sim y_j \in \mathcal{E}} \hat{V}(y_i, y_j) + \sum_{\substack{y_k \sim y_i \\ y_k \sim y_i \notin \mathcal{E}}} \hat{V}(y_k, y_i) \qquad (15) \\
&= \sum_{\substack{y_j \\ y_i \sim y_j \in \mathcal{E}}} D(y_j) + \sum_{y_i \sim y_j \in \mathcal{E}} (\max(|D(y_i)|, |D(y_j)|) + \epsilon) + \sum_{\substack{y_k \sim y_i \\ y_k \sim y_i \notin \mathcal{E}}} \hat{V}(y_k, y_i) > 0.
\end{aligned}$$

As a result $\hat{E}(Z) < \hat{E}(Y)$. Moreover since $E(Z) = \hat{E}(Z)$ (using Lemma 3),

$$E(Z) < \hat{E}(Y) < \hat{E}(X^*) < E(X^*). \qquad (16)$$

As a result, $E(Z) < E(X^*)$, which contradicts (13). Consequently X^* is the global minimum of $\hat{E}(X)$.

□