The Binary Space Partitioning-Tree Process Supplementary Material

Xuhui Fan
School of Mathematics and Statistics
University of New South Wales
xuhui.fan@unsw.edu.au

Bin Li
School of Computer Science
Fudan University
libin@fudan.edu.cn

Scott A. Sisson
School of Mathematics and Statistics
University of New South Wales
scott.sisson@unsw.edu.au

A Justification on the accumulated cut cost

Let $D_{\square} = \sup_{x_1, x_2 \in \square}\{|x_1 - x_2|\}$ denotes the diameter of the polygon $\square$. $\forall\theta$, it is obviously that the length of the newly generated cut line $L(\theta, \boldsymbol{u})$ is smaller or equal to $D_{\square}$, i.e., $|L(\theta, \boldsymbol{u})| \leq D_{\square}$. Thus, we have the result for the sum of perimeters in the l-th partitioning result as:

$$\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)}) \leq PE(\square) + 2(l-1)D_{\square} \tag{1}$$

According to the Fatou's lemma, we get

$$\begin{aligned}
&\mathbb{E}\liminf_{l\to\infty} \frac{\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)})}{l} \\
\leq &\liminf_{l\to\infty} \mathbb{E}\frac{\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)})}{l} \\
\leq &\liminf_{l\to\infty} \frac{\mathbb{E}[PE(\square)] + 2(l-1)D_{\square}}{l} \\
< &\infty
\end{aligned} \tag{2}$$

which leads to $\liminf_{l\to\infty} \frac{\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)})}{l} < \infty$ almost surely. Since $\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)})$ is increasing for l almost surely, we get $\sum_{l=1}^{\infty}\left[\sum_{k=1}^{l} PE(\square_{\tau_l}^{(k)})\right]^{-1} = \infty$ almost surely.

B Mathematical formulation of the three-restrictions on the measure invariance

1. translation t: $\lambda_{\square}(C_{\square}^{\theta}) = \lambda_{t_{\boldsymbol{v}}\square} \circ t_{\boldsymbol{v}}(C_{\square}^{\theta})$, where $t_{\boldsymbol{v}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{v}, \forall \boldsymbol{v} \in \mathbb{R}^2$;

2. rotation r: $\lambda_{\square} = \lambda_{r_{\theta'}\square} \circ r_{\theta'}$, where $r_{\theta'}(\boldsymbol{x}) = \begin{bmatrix} \cos\theta' & -\sin\theta' \\ \sin\theta' & \cos\theta' \end{bmatrix} \cdot \boldsymbol{x}$ refers to rotate the point $\boldsymbol{x}$ in an angle of θ';

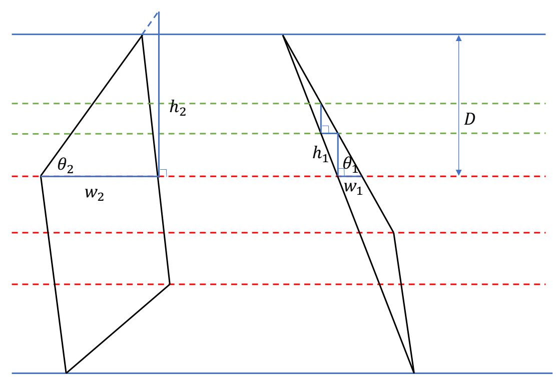

Figure 1: The design of a set of divisions in Lemma 1. Red dashed lines denote the "rough" divisions, while green dashed lines denote the "grained" divisions based on each consecutive rough divisions.

3. restriction ψ: $\lambda_{\square}(C_{\triangle}^{\theta}) = \lambda_{\psi_{\triangle}\square} \circ \psi_{\triangle}(C_{\triangle}^{\theta})$, where $\triangle \subseteq \square$ refers to a sub-domain of $\square$; $\psi_{\triangle}\square = \{\boldsymbol{x}|\boldsymbol{x} \in \triangle \subset \square\}$, and $C_{\triangle}^{\theta}$ refers to the set of cut lines for all the potential cuts crossing through $\triangle$.

C Proof of Proposition 1

Lemma 1. *Assume two convex polygons $\square_1$ and $\square_2$ have the same length on the line segment $\boldsymbol{l}(\theta)$. There exists a set of countable divisions passing through $\boldsymbol{l}(\theta)$ in the direction of $\theta + \frac{\pi}{2}$. Each line sub-segment of $\boldsymbol{l}(\theta)$ cut by consecutive divisions is covered by the intersection of the $\square_1, \square_2$, while $\square_1, \square_2$ can move in the direction of $\theta + \pi/2$.*

Proof. The set of divisions (all in the direction of $\theta + \frac{\pi}{2}$) can be designed into two stages.

Stage 1, a set of "rough" divisions that pass through each vertices of the two polygons.

Stage 2, sets of "grained" divisions based on the each consecutive rough divisions. Let D denotes the distance between two selected consecutive divisions, w_1 (w_2) denotes the maximum width of polygon $\square_1$ ($\square_2$) between these two rough divisions and θ_1 (θ_2) is the smallest angle between the edge of polygon $\square_1$ ($\square_2$) and division.

We proceed the grained division in the following way. In the case of $\min\{w_1 \tan\theta_1, w_2 \tan\theta_2\} \geq D$, there is no need to do further grained division; otherwise, let $(\theta^*, d^*) = \arg_{\theta,d} \min\{d_1 \cdot \tan\theta_1, d_2 \cdot \tan\theta_2\}$, the first grained division is placed in the position of $d^* \tan\theta^*$ (see h_1 in Figure 1). The second grained division would design based on the first one and proceed in a similar way.

Given the condition of $d{\cdot}\tan\theta < D$, we can do the partition at the positions of:

$$\left\{ d \cdot \tan\theta \cdot \left(1 - \frac{d \cdot \tan\theta}{D}\right)^l \right\}_{l=0}^{\infty} \tag{3}$$

where $\sum_{l=0}^{\infty} d \cdot \tan\theta \cdot \left(1 - \frac{d\cdot\tan\theta}{D}\right)^l = D$.

□

Proposition 1. *The family of partition probability measure $\lambda_\square(C_\square^\theta)$ keeps invariant under the operations of translation, rotation and restriction if and only if we have a constant C such that $\lambda_\square(C_\square^\theta) = C \cdot |\boldsymbol{l}(\theta)|, \forall C \in \mathbb{R}^+$.*

Proof. The reverse case is fully discussed as in the main part of the paper.

On the other hand, assume we have two sets of cut lines $C_{\square_1}^{\theta_1}, C_{\square_2}^{\theta}$ with $|\boldsymbol{l}_{\square_1}(\theta_1)| = |\boldsymbol{l}_{\square_2}(\theta)|$.

Given that the measure $\lambda_\square(C_\square^\theta)$ is invariant under the operations of translation, rotation and restriction, we need to prove the following equity:

$$\lambda_{\square_1}(C_{\square_1}^{\theta_1}) = \lambda_{\square_2}(C_{\square_2}^{\theta}) \tag{4}$$

To complete this, we first do rotation and translation operations on $\square_1$, which is $\square_1' := r_{\theta'} \circ t_{\boldsymbol{v}'} \circ \square_1$, in a way that $\square_1'$ and $\square_2$ project into the same image $\boldsymbol{l}(\theta)$.

Based on Lemma 1, we divide $\square_1'$ and $\square_2$ into countable parts, where the intersection of these pair parts projects to the same images. That is:

$$\square_1' = \cup_k \square_1^{',(k)}, \square_2 = \cup_k \square_2^{(k)} \tag{5}$$

$$Y^{(k)} = \square_1^{',(k)} \cap \square_2^{(k)}, \boldsymbol{l}^{(k)}(\theta) \in Y^{(k)}, \forall k \in N \tag{6}$$

$$\begin{aligned} C_{\square_1^{',(k)}}^\theta &= \{L(\theta, \boldsymbol{u}) \text{ crossing } \square_1^{',(k)} | \theta \text{ is fixed}, \boldsymbol{u} \text{ lies on } \boldsymbol{l}^{(k)}(\theta)\} \\ C_{\square_2^{(k)}}^\theta &= \{L(\theta, \boldsymbol{u}) \text{ crossing } \square_2^{(k)} | \theta \text{ is fixed}, , \boldsymbol{u} \text{ lies on } \boldsymbol{l}^{(k)}(\theta)\} \end{aligned} \tag{7}$$

The additivity of measures indicates that:

$$\begin{aligned} \lambda_{\square_1}(C_{\square_1}^\theta) &= \sum_k \lambda_{\square_1^{',(k)}}(C_{\square_1^{',(k)}}^\theta) \\ \lambda_{\square_2}(C_{\square_2}^\theta) &= \sum_k \lambda_{\square_2^{(k)}}(C_{\square_2^{(k)}}^\theta) \end{aligned}$$

Eq. (4) is correct if we can prove the following

$$\lambda_{\square_1^{',(k)}}(C_{\square_1^{',(k)}}^\theta) = \lambda_{\square_2^{(k)}}(C_{\square_2^{(k)}}^\theta), \forall k \in N \tag{8}$$

From the measure invariance under rotation and translation, we get $\lambda_{\square_1}(C_{\square_1}^\theta) = \lambda_{t_{\boldsymbol{v}}(r_{\theta'}(\square_1))}(t_{\boldsymbol{v}}(r_{\theta'}(C_{\square_1}^\theta)))$.

We also get

$$\Pi_{Y^{(k)}}\pi = \{Y^{(k)}\} = \Pi_{Y^{(k)}} t^{(k)}(\rho\pi) \tag{9}$$

$$\Pi_{Y^{(k)}} C_{\square_1^{',(k)}}^\theta = \Pi_{Y^{(k)}} C_{\square_2^{(k)}}^\theta \tag{10}$$

Thus, we get

$$\begin{aligned} \lambda_{t_{\boldsymbol{v}} r_{\theta'} \square_1^{',(k)}}(C_{\square_1^{',(k)}}^\theta) &\overset{restriction}{=} \lambda_{\Pi_{Y^{(k)}} t_{\boldsymbol{v}} r_{\theta'} \square_1^{',(k)}}(\Pi_{Y^{(k)}} C_{\square_1^{',(k)}}^\theta) \\ &\overset{Eq.(9)}{=} \lambda_{\Pi_{Y^{(k)}} \square_1^{',(k)}}(\Pi_{Y^{(k)}} C_{\square_1^{',(k)}}^\theta) \\ &\overset{Eq.(10)}{=} \lambda_{\Pi_{Y^{(k)}} \square_2^{(k)}}(\Pi_{Y^{(k)}} C_{\square_2^{(k)}}^\theta) \\ &\overset{restriction}{=} \lambda_{\square_2^{(k)}}(C_{\square_2^{(k)}}^\theta) \end{aligned} \tag{11}$$

□

D Proof of Proposition 2

Our partition would result in convex polygon. We have the integration results for the convex polygon.

Lemma 2. *The integration of the intersection line in a triangle $\triangle$ over $[0, \pi]$ equals to the triangle's perimeter.*

Proof. We first consider the acute triangle (Top row of *Figure 2*) case.

Let $\{l_1, l_2, l_3\}$ being the lengths of the triangle $\triangle$'s edges and $\{\angle BAC, \angle ABC, \angle ACB\}$ being the corresponding angles. According to the law of sines, we have

$$l_0 = \frac{l_1}{\sin\angle BAC} = \frac{l_2}{\sin\angle ABC} = \frac{l_3}{\sin\angle ACB} \tag{12}$$

where we use l_0 to denote the ratio between the length and its corresponding angle.

W.l.o.g., we are cutting the block in the direction within $\angle ABC$. The projection scalar of l_2 is calculated as $|BD| =$

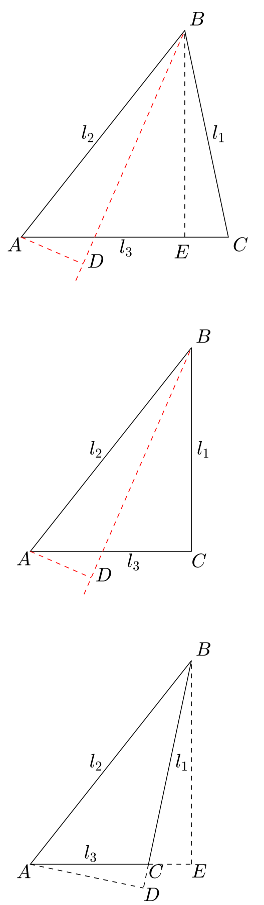

Figure 2: Top: Acute Triangle; Middle: Right Triangle; Bottom: Obtuse Triangle.

$l_2 \cos\theta(\theta = \angle ABD)$. While θ is ranging from 0 to $\angle ABE$, the integration of BD is

$$\begin{aligned}\int_0^{\angle ABE} |BD| d\theta = & \int_0^{\angle ABE} l_2 \cos\theta d\theta \\ = & l_2 \sin\theta|_0^{\angle ABE} = l_2\cos(\angle BAC)\end{aligned} \tag{13}$$

By using the similar routines, we can get the integration of all the projection lines I as:

$$\begin{aligned}I = & l_2\cos\angle BAC + l_2\cos\angle ABC + l_1\cos\angle ACB \\ & + l_1\cos\angle ABC + l_3\cos\angle BAC + l_3\cos\angle ACB \\ = & l_0\sin\angle ACB\cos\angle BAC + l_0\sin\angle ACB\cos\angle ABC \\ & + l_0\sin\angle BAC\cos\angle ACB + l_0\sin\angle BAC\cos\angle ABC \\ & + l_0\sin\angle ABC\cos\angle BAC + l_0\sin\angle ABC\cos\angle BAC \\ = & l_0\sin(\angle ACB + \angle ABC) + l_0\sin(\angle BAC + \angle ACB) \\ & + l_0\sin(\angle ABC + \angle BAC) \\ = & l_0\sin\angle BAC + l_0\sin\angle ABC + l_0\sin\angle ACB \\ = & l_1 + l_2 + l_3 = PE(\triangle)\end{aligned} \tag{14}$$

Here the 2$^{\text{nd}}$ equation holds due to the law of Sines.

The case of right triangle (Middle row of *Figure 2*) is straight forward.

We can get

$$\begin{aligned}I = & l_2\cos\angle BAC + l_2\cos\angle ABC \\ & + l_1\cos\angle ABC + l_3\cos\angle BAC \\ = & l_1 + l_2 + l_3 = PE(\triangle)\end{aligned} \tag{15}$$

On the case of obtuse triangle (Bottom row of *Figure 2*)

$$\begin{aligned}I = & l_2\cos\angle BAD - l_3\cos\angle CAD + l_2\cos\angle ABE \\ & - l_1\cos\angle CBE + l_3\cos\angle BAC + l_1\cos\angle ABC \\ = & l_1 + l_2 + l_3 = PE(\triangle)\end{aligned} \tag{16}$$

□

Lemma 3. *The integration of the length of the block's projected image in the direction of θ over $(0, \pi]$ equals to the perimeter of the block, which is $\int_0^\pi |\boldsymbol{l}(\theta)| d\theta = PE(\square)$.*

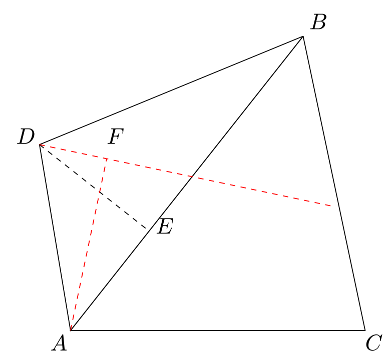

Figure 3: From convex polygon with $n-1$ vertices to convex polygon with n vertices.

Proof. Convex polygon with n vertices can be divided into $n-2$ triangles. Since we have the result for the case of triangles, mathematical induction is used to get the conclusion for any convex polygons.

Assume we have the result for convex polygon with $n-1$ vertices, the additional part for its transformation to convex polygon with n vertices is the triangle $\triangle ABD$. Correspondingly, the increase in the scalar projection is composed of two parts:

$$\begin{aligned}L^1_{\text{increase}} = & \int_0^{\angle DAB} |AD| \sin\theta d\theta \\ = & |AD| - |AD|\cos\angle DAB = |AD| - |AE|\end{aligned} \tag{17}$$

where $\theta = \angle DAF$ and L^1_{increase} refers to the integration of $|DF|$ in the angle of $\angle DAB$.

$$L^2_{\text{increase}} = \int_0^{\angle ABD} |BD| \sin\theta d\theta$$
$$= |BD| - |BD| \cos\angle ABD = |BD| - |BE| \quad (18)$$

Thus, the total add amount is $L_{\text{increase}} = |BD| + |AD| - (|AE| + |BE|) = |BD| + |AD| - |AB|$. This is exactly the increase of perimeter from the convex polygon with $n-1$ vertices to convex polygon with n vertices. Thus, we can get the result for all the convex polygons. □

Proposition 2 is a direct result of Lemma 3.

E Consistency

Some notations are firstly defined for convenient reference. We use $\Box$ and $\triangle$ to denote a domain and its subdomain, which is $\triangle \subseteq \Box$. M_τ and N_τ are individually defined as the BSP-Tree processes on $\Box$ and $\triangle$ respectively. The restriction is denoted as Π, in which we have $\Pi_\triangle M_\tau = N_\tau$. Also, we let $c(\Box)$ denote the measure over the block $\Box$, which is $c(\Box) = \int_\theta \omega(\theta)|\boldsymbol{l}_\Box(\theta)|d\theta$ and let $\mathcal{O}^\Box_n$ denote the partition after n-th cut on the convex polygon $\Box$.

Extending partition from $\triangle$ to $\Box$ For the BSP-Tree process N_τ, we let Z and $\{\sigma_l\}_{l\in\mathbb{N}}$ denotes the related Markov chain and the corresponding time stops. For $t \geq 0$, define m_t to be the index such that $t \in [\sigma_{m_t}, \sigma_{m_{t+1}}]$, $N_t = Z_{m_t}$.

To extend N_τ from $\triangle$ to $\Box$, let $\tau_0 = 0$ and $Y_0 = \Box$. For $n \in \mathbb{N}$, we define τ_{n+1} and Y_{n+1} inductively as:

$$\tau_{n+1} := \min\{\sigma_{m_{\tau_n}+1}, \tau_n + \frac{\xi_n}{c(Y_n) - c(Z_{m_{\tau_n}})}\} \quad (19)$$

where ξ_n is generated from the exponential distribution with mean 1.

$$Y_{n+1} = \begin{cases} lift_{Y_n,\triangle}(Z_{m_{\tau_n}}), & \tau_{n+1} = \sigma_{m_{\tau_n}+1}; \\ gencut_\triangle(Y_n), & \text{otherwise.} \end{cases} \quad (20)$$

where $lift_{Y_n,\triangle}(Z_{m_{\tau_n}})$ denotes extending the existing cut to the larger domain $\Box$ and $gencut_\triangle(Y_n)$ refers to the case that there will be a new cut generated in $\Box$ that does not cross into $\triangle$.

According to the results of Proposition V.16 of chapter VI in [1], the defined process $\{Y_n, \tau_n\}$ are well-defined.

Prove the correctness $\forall t > 0$, the waiting time for the next cut in X is:

$$\zeta_t = \tau_{n+1} - t \quad (21)$$

According to τ_{n+1}'s definition (Eq. (19)), ζ_t follows the exponential distribution with the rate being $c(Y_{n_t})$. What is more, the probability of the event $\tau_{n+1} = \sigma_{m_{\tau_n}+1}$ occurs with probability $c(\mathcal{O}^\triangle_n)/c(\mathcal{O}^\Box_n)$.

For the newly extended case $\{Y_n\}_n$, while Y_n crosses through $\triangle$, the probability measure on $C^\theta_\triangle$ is in proportion to $\omega(\theta)|\boldsymbol{l}_\triangle(\theta)|$ and $\boldsymbol{u}$ locates only on $\boldsymbol{l}_\triangle(\theta)$. Thus, we get

$$P = \frac{c(\mathcal{O}^\triangle_n)}{c(\mathcal{O}^\Box_n)} \cdot \frac{\omega(\theta)|\boldsymbol{l}_\triangle(\theta)|}{\int_\theta \omega(\theta)|\boldsymbol{l}_\triangle(\theta)|d\theta} \cdot \frac{1}{|\boldsymbol{l}_\triangle(\theta)|}$$
$$= \frac{\omega(\theta)}{c(\mathcal{O}^\Box_n)} \quad (22)$$

while Y_n does not cross through $\triangle$, , the probability measure on $C^\theta_{\Box\setminus\triangle}$ is in proportion to $\omega(\theta)(|\boldsymbol{l}_\Box(\theta)| - |\boldsymbol{l}_\triangle(\theta)|)$ and $\boldsymbol{u}$ locates only on $\boldsymbol{l}_{\Box\setminus\triangle}(\theta)$ (with the length $|\boldsymbol{l}_\Box(\theta)| - |\boldsymbol{l}_\triangle(\theta)|$). Thus, we get

$$P = \frac{c(\mathcal{O}^\Box_n) - c(\mathcal{O}^\triangle_n)}{c(\mathcal{O}^\Box_n)} \cdot \frac{\omega(\theta)(|\boldsymbol{l}_\Box(\theta)| - |\boldsymbol{l}_\triangle(\theta)|)}{\int_\theta \omega(\theta)(|\boldsymbol{l}_\Box(\theta)| - |\boldsymbol{l}_\triangle(\theta)|)d\theta}$$
$$\cdot \frac{1}{|\boldsymbol{l}_\Box(\theta)| - |\boldsymbol{l}_\triangle(\theta)|}$$
$$= \frac{\omega(\theta)}{c(\mathcal{O}^\Box_n)} \quad (23)$$

Eq. (22) and Eq. (23) show that the probability measure of Y_n equals to the one that directly generated in the domain of $\Box$. Therefore, the partition constructed by Eq. (19) and Eq. (20) is a realization of BSP-Tree process in $\Box$.

According the transfer theorem (Theorem V.13 of chapter VI in [1]), the partition distribution is consistent from $\Box$ to $\triangle$.

F MCMC for the BSP-RM

Algorithm 2 displays an MCMC solution for the BSP-RM.

Algorithm 2 MCMC for BSP-RM

Input: Training data X, Budget τ, Number of particles C

Output: A realization of the BSP-Tree process; coordinates $\{(\xi_i, \eta_i)\}_{i=1}^n$ of X

1: Initialize the partition and nodes' coordinates
2: **for** $t = 1 : T$ **do**
3: Use C-SMC algorithm to update the partition structure, according to Algorithm **1**;
4: Update nodes' coordinates $\{\xi_i, \eta_i\}_{i=1}^n$ according to Eq. (24).
5: **end for**

F.1 Updating nodes' coordinates $(\xi_i, \eta_i)_{i=1}^n$

(ξ_i, η_i)'s updating is implemented through the Metropolis-Hastings algorithm. We propose the new values of ξ_i, η_i

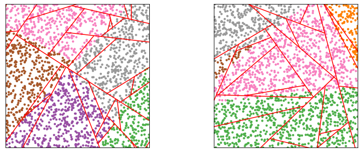

Figure 4: Toy Data Partition Visualization (Case 1).

with the uniform distribution in $[0, 1]$ and the acceptance ratio $\min(1, \alpha)$ is as follows:

$$\begin{aligned}\alpha(\xi_i, \xi_i^0) &= \frac{\prod_{j'=1}^{n} P(e_{ij'}|\xi_i, \xi_{\setminus i}, \eta_{j'}, \theta)}{\prod_{j'=1}^{n} P(e_{ij'}|\xi_i^0, \xi_{\setminus i}, \eta_{j'}, \theta)};\\ \alpha(\eta_j, \eta_j^0) &= \frac{\prod_{i'=1}^{n} P(e_{i'j}|\eta_j, \eta_{\setminus j}, \xi_{i'}, \theta)}{\prod_{i'=1}^{n} P(e_{i'j}|\eta_j^0, \eta_{\setminus j}, \xi_{i'}, \theta)}\end{aligned} \tag{24}$$

G Visualization of Case 1

Figure 4 shows the visualization of Case 1.

References

[1] Daniel M. Roy. *Computability, Inference and Modeling in Probabilistic Programming*. PhD thesis, MIT, 2011.