A process convergence result for partial match queries in random quadtrees

Henning Sulzbach
J. W. Goethe-Universität Frankfurt a. M.

Mini-Workshop: Random Trees, Information and Algorithms
Oberwolfach, April 28, 2011

joint work with Nicolas Broutin and Ralph Neininger
Partial match queries

A classical combinatorial problem is to perform a search in a multidimensional database where the record to be retrieved is either fully or partially specified. The latter is called a Partial Match Query.

\( n \)-dim. domain: \( S = S_1 \times \cdots \times S_n \)

Problem: For a fixed query \( q = (q_1, \ldots, q_n) \) with \( q_i \in S_i \cup \{\ast\} \) and a set of data \( S' \subseteq S \) find all elements \( s = (s_1, \ldots, s_n) \in S' \) such that

\[ s_i = q_i, \quad \text{if} \quad q_i \neq \ast. \]
Partial match queries

Comparison-based structures:

- Quadtree (Finkel and Bentley ’74),
- K-d-tree (Bentley ’75)

Several variants are known in the literature, among them
K-d-trees (Cunto, Lau, Flajolet ’89), random relaxed K-d-trees
(Duch, Estivill-Casto, Martinez ’98) and squarish K-d-tress
(Devroye, Jabbour, Zamora-Cura ’99).

Digital structures:

- K-d-tries (Rivest ’76)
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: \( S_i = [0, 1] \) for all \( i \).

Quadtree: \( d = 2 \)
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: \( S_i = [0, 1] \) for all \( i \).

Quadtree: \( d = 2 \)

\[
\begin{bmatrix}
0.73 & 0.17 \\
0.1 & 0.86
\end{bmatrix}
\]
Comparison-based structures

Model: \( S_i = [0, 1] \) for all \( i \).

Quadtree: \( d = 2 \)
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: \( S_i = [0, 1] \) for all \( i \).

Quadtree: \( d = 2 \)

0.56, 0.3
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: $S_i = [0, 1]$ for all $i$.

Quadtree: $d = 2$
Comparison-based structures

Model: \( S_i = [0, 1] \) for all \( i \).

Quadtree: \( d = 2 \)
Comparison-based structures

2-d-tree:
Comparison-based structures

2-d-tree:  

\[0.69, 0.73\]
Comparison-based structures

2-d-tree:
Comparison-based structures

2-d-tree:

0.69
0.73

0.17, 0.46
Comparison-based structures

2-d-tree:

```
0.69
0.73

0.17
0.46
```
Comparison-based structures

2-d-tree:

![2-d-tree diagram]

0.26, 0.86
Comparison-based structures

2-d-tree:

```
0.69
0.73
0.17
0.46
0.26
0.86
```

A process convergence result for partial match queries in random quadtrees.
Comparison-based structures

2-d-tree:

```
0.69
0.73

0.17
0.46

0.26
0.86

0.8, 0.72
```
Comparison-based structures

2-d-tree:

```
    0.69
   /   
0.17  0.8
 /     /  
0.26  0.46  0.72
```

A process convergence result for partial match queries in random quadtrees

Henning Sulzbach
J. W. Goethe-Universität Frankfurt a. M.
Comparison-based structures

2-d-tree:

```
0.69 0.73
0.17 0.46
0.26 0.86
0.8 0.72
```

```
0.54, 0.1
```
Comparison-based structures

2-d-tree:
Comparison-based structures

2-d-tree:

0.19, 0.9
Comparison-based structures

2-d-tree:
Comparison-based structures

2-d-tree:

```
0.69 0.73
0.17 0.46
0.54 0.1
0.26 0.86
```

```
0.8 0.72
0.19 0.9
```

```
0.6, 0.36
```
Comparison-based structures

2-d-tree:
Partial Match query - Quadtrees

Query: \( q = \{s, *\} \), \( s = 0.2 \)
Partial Match query - Quadtree

Query: $q = \{s, *\}$, $s = 0.2$
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, \hspace{1cm} $s = 0.2$
Partial Match query - Quadtree

Query: \( q = \{s, \ast\}, \quad s = 0.2 \)
Partial Match query - Quadtree

Query: \( q = \{s, *\} \), \( s = 0.2 \)
Partial Match query - Quadtree

Query: $q = \{s, *\}$, \hspace{1cm} s = 0.2

$\text{s}$
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, \hspace{1cm} $s = 0.2$

\begin{tikzpicture}
  \node[shape=circle,fill=green] (0) at (0,0) {0.73
  \begin{tikzpicture}
    \node[shape=circle,fill=green] (1) at (-1,0) {0.17};
    \node[shape=rectangle,draw] (2) at (-2,0) {};
    \node[shape=rectangle,draw] (3) at (-1,1) {};
    \node[shape=rectangle,draw] (4) at (-1,-1) {};
    \node[shape=rectangle,draw] (5) at (0,0) {};
  \end{tikzpicture}
\end{tikzpicture}

\end{tikzpicture}
Partial Match query - Quadtree

Query: \( q = \{s, *\} \), \hspace{1cm} s = 0.2
Partial Match query - Quadtree

Query: \( q = \{s, *\}, \quad s = 0.2 \)
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, $s = 0.2$
Partial Match query - Quadtree

Query: $q = \{s, *\}$, \hspace{1cm} s = 0.2

Henning Sulzbach
J. W. Goethe-Universität Frankfurt a. M.

A process convergence result for partial match queries in random quadtrees
Partial Match query - Quadtree

Query: \( q = \{s, \ast\}, \quad s = 0.2 \)
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, $s = 0.2$
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, \hspace{1cm} s = 0.2
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, \hspace{1cm} $s = 0.2$
Partial Match query - Quadtrees

Query: $q = \{s, *\}, \quad s = 0.2$
Partial Match query - Quadtree

Query: $q = \{s, \ast\}$, \hspace{1cm} s = 0.2

Henning Sulzbach J. W. Goethe-Universität Frankfurt a. M.

A process convergence result for partial match queries in rand...
Partial Match query - Quadtree

Query: $q = \{s, *\}$, \hspace{1cm} s = 0.2
Partial Match Queries - Quadtrees

Observation:

Performing a partial match query with \( q = \{s, \ast\} \), a node is visited if and only if it is inserted in a subregion that intersects the line \( \{x = s\} \).

This is equivalent to an intersection of its horizontal line and \( \{x = s\} \).
Partial Match query - 2-d-tree

Query: \( q = \{s, *\} \), \( s = 0.55 \)
Partial Match query - 2-d-tree

Query: $q = \{s, \ast\}$, \hspace{1cm} s = 0.55

```
0.69
0.73

0.17
0.46

0.54
0.1

0.6
0.36

0.26
0.86

0.8
0.72

0.17
0.46

0.26
0.86

0.6
0.36

0.54
0.1

0.19
0.9

0.1
0.6
0.36
0.19
0.9
```
Partial Match query - 2-d-tree

Query: \( q = \{s, *\}, \quad s = 0.55 \)
Partial Match query - 2-d-tree

Query: $q = \{s, \ast\}$, $s = 0.55$
Partial Match query - 2-d-tree

Query: $q = \{ s, * \}$, \hspace{1cm} s = 0.55
Partial Match query - 2-d-tree

Query: $q = \{s, *\}, \quad s = 0.55$
Partial Match query - 2-d-tree

Query: $q = \{*, s\}$, \hspace{1cm} $s = 0.5$
Partial Match query - 2-d-tree

Query: \( q = \{*, s\}, \quad s = 0.5 \)
Partial Match query - 2-d-tree

Query: $q = \{\ast, s\}, \quad s = 0.5$
Partial Match query - 2-d-tree

Query: \( q = \{ *, s \} \), \( s = 0.5 \)
Partial Match query - 2-d-tree

Query: $q = \{\ast, s\}, \quad s = 0.5$
Partial Match query - 2-d-tree

Query: \( q = \{\ast, s\}, \quad s = 0.5 \)
Probabilistic model

For the analysis of the complexity of a partial match query we always assume the coordinates of elements in the database $S'$ to be *independent* and *uniform* on $[0, 1]$.

$C_n(s)$: number of nodes visited by a partial match query for $q = \{s, \ast\}$ in a random two-dimensional quadtree of size $n$.  

Henning Sulzbach
J. W. Goethe-Universität Frankfurt a. M.
A process convergence result for partial match queries in random
Simulation - Quadtree

\[ n = 100 \]
Simulation - Quadtree

\[ n = 200 \]
Two remarks

In a random quadtree of size $n$

$$\mathbb{P}({\text{First subtree is empty}}) = \mathbb{E}[(1 - U V)^n] = \frac{M_{n,2}}{n},$$

where $M_{n,2}$ is the number of maxima.

Forgetting the labels,

- a 2d-tree is a random Binary Search tree,
- a Quadtree is not a random 4-ary recursive tree.
First analysis on $C_n(s)$

$X$ uniform on $[0, 1]$, independent of the quadtree.

Claim (Bentley, Stanat ’75)

For $n \to \infty$, it holds $\mathbb{E}[C_n(X)] \asymp n^{1/2}$.

Heuristic: Let $l_1^{(n)}, l_2^{(n)}, l_3^{(n)}, l_4^{(n)}$ be the subtree sizes.
Assuming
\[ l_r^{(n)} \approx n/4, \quad r = 1, \ldots, 4 \]
leads to
\[ \mathbb{E}[C_n(X)] \approx 1 + 2\mathbb{E}[C_{n/4}(X)] \]
which implies
\[ \mathbb{E}[C_n(X)] \asymp n^{1/2}. \]
The claim is wrong, a quadtree is too far from being perfect.
Asymptotic results on $C_n(X)$

Theorem (Flajolet, Gonnet, Puech, Robson ’93)

For $n \to \infty$, it holds

$$\mathbb{E}[C_n(X)] \sim \kappa n^\beta$$

with

$$\kappa = \frac{\Gamma(2\beta + 2)}{2\Gamma^3(\beta + 1)} \approx 1.59, \quad \beta = \frac{\sqrt{17} - 3}{2} \approx 0.56.$$ 

Proof [Sketch]: Given the first key $(U, V)$ it holds

$$(l_1^{(n)}, l_2^{(n)}, l_3^{(n)}, l_4^{(n)}) \overset{d}{=} \text{Mult}(n-1; UV, U(1-V), (1-U)V, (1-U)(1-V)).$$
Asymptotic results on $C_n(X)$

This gives

$$\mathbb{E}[C_n(X)] = 1 + \frac{4}{n(n+1)} \sum_{k=0}^{n-1} (n-k) \mathbb{E}[C_k(X)].$$

This implies a differential equation for

$$Q(z) = \sum_{n \geq 0} \mathbb{E}[C_n(X)] z^n$$

that gives

$$Q(z) = \frac{2 F_1[-\beta - 1, -\beta; 2; z]}{(1-z)^{\beta+1}} - \frac{1}{1-z}.$$

The statement then follows by Singularity analysis.
Further asymptotic results on the mean

Chern, Hwang ’05: \( \mathbb{E}[C_n(X)] = \kappa n^\beta - 1 + O(n^{-\beta}). \)

Higher dimensions:

- order - Flajolet, Gonnet, Puech, Robson ’93,
- leading constants, lower order terms - Chern, Hwang ’05.

Theorem (Curien, Joseph ’11)

For fixed \( s \in [0, 1] \) and \( n \to \infty \), it holds

\[
\mathbb{E}C_n(s) \sim K_1 n^\beta (s(1 - s))^{\beta/2},
\]

where

\[
K_1 \int_0^1 (s(1 - s))^{\beta/2} ds = \kappa.
\]
Recursive decomposition

For $X_n = C_n(X)$ the decomposition

$$X_n \overset{d}{=} 1 + 1\{X < U\} \left( X_{I_1}^{(1)} + X_{I_2}^{(2)} \right) + 1\{X \geq U\} \left( X_{I_3}^{(3)} + X_{I_4}^{(4)} \right),$$

where $(X_n^{(1)}, X_n^{(2)}, X_n^{(3)}, X_n^{(4)})$ are ind. copies of $(X_n)$, ind. of $(U, V, I_1^{(n)}, I_2^{(n)}, I_3^{(n)}, I_4^{(n)})$ is wrong!

There a dependencies between $X^{(1)}$ and $X^{(2)}$ resp. $X^{(3)}$ and $X^{(4)}$.

Still

$$\mathbb{E}[X_n] = 1 + 2 \left( \mathbb{E} \left[ 1\{X < U\} X_{I_1}^{(1)} \right] + \mathbb{E} \left[ 1\{X \geq U\} X_{I_3}^{(3)} \right] \right).$$
Recursive decomposition

This allows us to compute $\beta$:

$$E[X_n] = 1 + 2 \left( E \left[ 1\{X<U\} X_{l_1(n)}^{(1)} \right] + E \left[ 1\{X\geq U\} X_{l_3(n)}^{(3)} \right] \right)$$

$$= 1 + 2E[X_{L_n}]$$

with $L_n \overset{d}{=} \text{Bin}(n - 1, \sqrt{UV})$. Scaling gives

$$n^{-\gamma}E[X_n] \sim 2E \left[ \left( \frac{L_n}{n} \right)^{\gamma} \frac{X_{L_n}}{L^n} \right].$$

Hence $1 = 2E[(\sqrt{UV})^\gamma] \Rightarrow \gamma = \beta$. 
Behaviour at the edge

Easier case $s = 0$:

$$C_n(0) \overset{d}{=} 1 + C_{l_1^{(n)}}^{(1)}(0) + C_{l_2^{(n)}}^{(2)}(0),$$

where $(C_n^{(1)})$, $(C_n^{(2)})$ are ind. copies of $(C_n)$, ind. of $(U, V, I_1^{(n)}, I_2^{(n)})$.

Mean: $\mathbb{E}[C_n(0)] \sim K_0 n^{\sqrt{2} - 1}$ (Flajolet, Gonnet, Puech, Robson ’93)

Theorem (Curien, Joseph ’11)

For $n \to \infty$, it holds

$$K_0^{-1} n^{1-\sqrt{2}} C_n(0) \overset{d}{\to} Z_0.$$
Behaviour at the edge

Theorem (Curien, Joseph ’11)

For $n \to \infty$, it holds

$$K_0^{-1} n^{1-\sqrt{2}} C_n(0) \xrightarrow{d} Z_0,$$

where $Z_0$ is the unique solution of

$$Z_0 \overset{d}{=} (UV)^{\sqrt{2}-1} Z_0^{(1)} + (U(1 - V))^{\sqrt{2}-1} Z_0^{(2)}$$

with unit mean and finite variance, where $Z_0^{(1)}, Z_0^{(2)}$ are ind. copies of $Z_0$, ind. of $(U, V)$.

Observe: Contraction in $\zeta_2$. 

Recursive decomposition

General case:

\[ C_n(s) \overset{d}{=} 1 + 1_{\{s < U\}} \left( C^{(1)}_{I_1^{(n)}} \left( \frac{S}{U} \right) + C^{(2)}_{I_2^{(n)}} \left( \frac{S}{U} \right) \right) \]

\[ + 1_{\{s \geq U\}} \left( C^{(3)}_{I_3^{(n)}} \left( \frac{s - U}{1 - U} \right) + C^{(4)}_{I_4^{(n)}} \left( \frac{s - U}{1 - U} \right) \right), \]

for fixed \( s \in [0, 1] \) and on the level of cadlag functions on \([0, 1] ,\)

where \((C^{(1)}_{n}), (C^{(2)}_{n}), (C^{(3)}_{n}), (C^{(4)}_{n})\) are ind. copies of \((C_{n}),\) ind. of \((U, V, I_1^{(n)}, I_2^{(n)}, I_3^{(n)}, I_4^{(n)})\).
Recursive decomposition

Scaling gives

$$\frac{C_n(s)}{n^\beta} \overset{d}{=} n^{-\beta} + 1 \{ s < U \} \left( \left( \frac{I_1^{(n)}}{n} \right)^\beta C_1^{(1)} \left( \frac{s}{U} \right) + \left( \frac{I_2^{(n)}}{n} \right)^\beta C_2^{(2)} \left( \frac{s}{U} \right) \right)$$

$$+ 1 \{ s \geq U \} \left( \left( \frac{I_3^{(n)}}{n} \right)^\beta C_3^{(3)} \left( \frac{s-U}{1-U} \right) + \left( \frac{I_4^{(n)}}{n} \right)^\beta C_4^{(4)} \left( \frac{s-U}{1-U} \right) \right).$$
Fixed-point equation

Assuming $n^{-\beta} C_n(s) \to Z(s)$ uniformly in $s \in [0, 1]$ for $n \to \infty$, suggests that $Z$ satisfies

$$Z(s) \overset{d}{=} 1_{\{s < U\}} (U V)^\beta Z^{(1)} \left( \frac{s}{U} \right) + (U(1 - V))^\beta Z^{(2)} \left( \frac{s}{U} \right)$$

$$+ 1_{\{s \geq U\}} ((1 - U)V)^\beta Z^{(3)} \left( \frac{s - U}{1 - U} \right)$$

$$+ 1_{\{s \geq U\}} ((1 - U)(1 - V))^\beta Z^{(4)} \left( \frac{s - U}{1 - U} \right), \quad (1)$$

where $Z^{(1)}, Z^{(2)}, Z^{(3)}, Z^{(4)}$ are ind. copies of $Z$, ind. of $(U, V)$. 
Results

Proposition (Existence)

There exists a unique random continuous process $Z(s)$ with
$\mathbb{E}[Z(s)] = (s(1-s))^{\beta/2}$, finite absolute moments of all orders such
that $Z(s)$ satisfies (1). Additionally,

$$\text{Var}(Z(s)) = K_2(s(1-s))^\beta, \quad K_2 = \frac{2(2\beta + 1)}{3(1 - \beta)} B(\beta + 1, \beta + 1) - 1.$$  

If $X$ is uniform on $[0, 1]$ and independent of $Z$, then

$$\text{Var}(Z(X)) = \frac{2(2\beta + 1)}{3(1 - \beta)} B^2(\beta + 1, \beta + 1) - B^2 \left( \frac{\beta}{2} + 1, \frac{\beta}{2} + 1 \right).$$
Theorem

Let $Z$ as in the previous Proposition. Then

$$\left( \frac{C_n(s)}{K_1 n^\beta} \right)_{s \in [0,1]} \rightarrow (Z(s))_{s \in [0,1]}, \quad n \rightarrow \infty,$$

in distribution in $(\mathcal{D}[0,1], \| \cdot \|_\infty)$, the space of cadlag functions on the unit interval endowed with the supremum norm. Moreover

$$n^{-\beta} \mathbb{E}[C_n(s)] \rightarrow K_1(s(1 - s))^{\beta/2}$$

and

$$n^{-2\beta} \text{Var}(C_n(s)) \rightarrow K_1^2 K_2(s(1 - s))^\beta.$$
Theorem (continuation)

If $X$ is uniformly distributed on $[0, 1]$, independent of $(C_n)$ and $Z$, then

$$\frac{C_n(X)}{K_1 n^\beta} \xrightarrow{d} Z(X)$$

and

$$\mathbb{E} \left[ \frac{C_n(X)}{K_1 n^\beta} \right] \to \mathbb{E}[Z(X)], \quad \mathbb{E} \left[ \left( \frac{C_n(X)}{K_1 n^\beta} \right)^2 \right] \to \mathbb{E}[Z(X)^2].$$

In particular

$$\text{Var} (C_n(X)) \sim K_4 n^{2\beta},$$

where

$$K_4 := K_1^2 \text{Var} (Z(X)) \approx 0.447363034.$$
Simulations

A process convergence result for partial match queries in random quadtrees

Henning Sulzbach J. W. Goethe-Universität Frankfurt a. M.
Simulations

A process convergence result for partial match queries in random quadtrees

Henning Sulzbach  J. W. Goethe-Universität Frankfurt a. M.
Simulations

n = 500
2d-tree - Asymptotics on the mean

By Flajolet, Puech ’86 and Chern, Hwang ’06

\[ \mathbb{E}[\bar{C}_n(X, \ast)] = \bar{\kappa}_1 n^\beta - 2 + O(n^{\beta-1}) \]

and

\[ \mathbb{E}[\bar{C}_n^{(2d)}(\ast, X)] = \bar{\kappa}_2 n^\beta - 3 + O(n^{\beta-1}), \]

where

\[ \bar{\kappa}_1 = \frac{13(3 - 5\beta)}{4} \kappa \approx 1.99, \quad \bar{\kappa}_2 = \frac{13(2\beta - 1)}{2} \kappa \approx 2.55. \]
Main Result - 2d-tree

Theorem

For \( \bar{K}_1 = \kappa_1 / \left( \int_0^1 (s(1 - s))^{\beta/2} \right) \),

\[
\left( \frac{\bar{C}_n(s, \ast)}{\bar{K}_1 n^\beta} \right)_{s \in [0, 1]} \rightarrow (\bar{Z}_1(s))_{s \in [0, 1]}, \quad n \rightarrow \infty,
\]

in distribution in \((\mathcal{D}[0, 1], \| \cdot \|_\infty)\), where \( \bar{Z}_1(s) \) is continuous and

\[
\bar{Z}_1(s) \overset{d}{=} 1_{\{s < U\}} \left( (UV_1)^\beta \bar{Z}_1^{(1)} \left( \frac{s}{U} \right) + (U(1 - V_1))^{\beta} \bar{Z}_1^{(2)} \left( \frac{s}{U} \right) \right)
+ 1_{\{s \geq U\}} \left( (1 - U)V_2 \right)^\beta \bar{Z}_1^{(3)} \left( \frac{s - U}{1 - U} \right)
+ 1_{\{s \geq U\}} \left( (1 - U)(1 - V_2) \right)^\beta \bar{Z}_1^{(4)} \left( \frac{s - U}{1 - U} \right).
\]
Main Result - 2d-tree

Theorem

For $\bar{K}_2 = \bar{\kappa}_2 / \left( \int_0^1 (s(1 - s))^{\beta/2} \right)$

$$
\left( \frac{C_n^{(2d)}(\ast, s)}{\bar{K}_2 n^{\beta}} \right)_{s \in [0,1]} \to (\bar{Z}_2(s))_{s \in [0,1]}, \quad n \to \infty,
$$

in distribution in $(\mathcal{D}[0, 1], \| \cdot \|_{\infty})$, where $\bar{Z}_2(s)$ is continuous and solves a similar fixed-point equation, namely
\[\bar{Z}_2(s) \overset{d}{=} 1_{\{s < u_1\}}(u_1v)^\beta \bar{Z}_2^{(1)} \left( \frac{s}{u_1} \right) \]

\[+ 1_{\{s < u_2\}}(u_2(1 - v))^\beta \bar{Z}_2^{(2)} \left( \frac{s}{u_2} \right) \]

\[+ 1_{\{s \geq u_1\}}((1 - u_1)v)^\beta \bar{Z}_2^{(3)} \left( \frac{s - u_1}{1 - u_1} \right) \]

\[+ 1_{\{s \geq u_2\}}((1 - u_2)(1 - v))^\beta \bar{Z}_2^{(4)} \left( \frac{s - u_2}{1 - u_2} \right). \]

Results for the variance can be obtained analogously.
Proof - Contraction method 😊

$$\mathbb{E}[(UV)^{2\beta}] < \frac{1}{4}$$

gives hope that contraction method with respect to $\zeta_2$ metric works.

Tasks:

- $s > 1 \Rightarrow$ exact scaling required. Therefore, necessary for the coefficient to converge, is

$$\sup_{s \in [0,1]} |n^{-\beta} \mathbb{E}[C_n(s)] - K_1(s(1 - s))^{\beta/2}| \to 0$$
Deducing weak convergence needs a rate on $\zeta_s$ convergence which comes from the rate for the coefficients. This follows from

$$\sup_{s \in [0,1]} |n^{-\beta} \mathbb{E}[C_n(s)] - K_1(s(1 - s))^{\beta/2}| = O(n^{-\varepsilon}).$$

for some $\varepsilon > 0$.

jump points of $(C_n(s))_{s \in [0,1]}$ are random, but their distance is large.
Proof - Convergence of mean

Idea: Continuous-time model $P_t(s)$

Points can be taken from an PPP on $[0,1]^2 \times [0,\infty)$. Let $\tau_1$ be the arriving time of the first point. Then

$$
\mathbb{E}[P_t(s)] = \mathbb{P}(t \geq \tau_1) + 2\mathbb{E}[\tilde{P}_{\text{Leb}}(Q_1(s))(t-\tau_1(s))(\xi_1(s))],
$$

where $(\tilde{P}_t)$ is an independent copy of $(P_t)$.
Iterating

\[ \text{Iterating} \]

\[ (U,V) \]

\[ Q \]

\[ S \]

\[ (U,V) \]

\[ Q_1 \]
Iterating

\[ (U,V) \]

\[ s \]

Henning Sulzbach  
J. W. Goethe-Universität Frankfurt a. M.

A process convergence result for partial match queries in random quadtrees
Iterating

\[ Q_2 \]

\((U, V)\)

\(s\)
Iterating

\[ \text{(U,V)} \]

\[ S \]

Henning Sulzbach  
J. W. Goethe-Universität Frankfurt a. M.

A process convergence result for partial match queries in random
Iterating
Iterating

\[(U, V)\]

\[s\]
Iterating

\[
\mathbb{E}[P_t(s)] = g_k(t) + 2^k \mathbb{E}[\tilde{P}_{\text{Leb}}(Q_k(s))(t-\tau_k(s))(\xi_k(s))],
\]

where \( g_k(t) \leq 2^k \) and \( \xi_k(s) \) is the position of \( s \) relative in \( Q_k(s) \).
Now,

\[
\frac{\mathbb{E}[P_t(s)]}{t^\beta} = t^{-\beta} g_k(t) + 2^k t^{-\beta} \mathbb{E}[\tilde{P}_{\text{Leb}(Q_k(s))(t-\tau_k(s))}(\zeta_k(s))] \\
= t^{-\beta} g_k(t) \\
+ 2^k \mathbb{E} \left[ (\text{Leb}(Q_k(s)) - t^{-1}\tau_k(s))^{\beta} \frac{\tilde{P}_{\text{Leb}(Q_k(s))(t-\tau_k(s))}(\zeta_k(s))}{(\text{Leb}(Q_k(s))(t - \tau_k(s)))^\beta} \right]
\]

Choosing \( k \) large allows to assume

\[
\zeta_k(s) \overset{d}{=} \text{unif}[0, 1].
\]

Then, known convergence results for \( \mathbb{E}[C_n(X)] \) can be applied.

Shape of the solution is obtained by a fixed-point argument.
Where does \((s(1-s))^{\beta/2}\) come from?

The relation on the previous slide already suggests that

\[ \kappa2^k\mathbb{E}[(\text{Leb}(Q_k(s)))^\beta] \rightarrow K_1(s(1-s))^{\beta/2}, \quad k \rightarrow \infty. \]

This can be made rigorous. Although being two-dimensional, only the \(x\)-coordinates of \(Q_k(s)\) is interesting, since

\[ \text{Leb}(Q_k(s)) = L_k(s) \prod_{\ell=1}^{k} X_\ell, \]

where \(X_1, X_2, \ldots\) is a sequence of ind. uniform \([0, 1]\) r.v., ind. of \(L_k(s)\).
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Here, $L_k(s)$ is the length of the interval on the $x$-axis covering $s$ after $k$ iterations.
Obviously,

\[ L_k(s) \overset{d}{=} \begin{cases} 1_{\{s<U\}} U L_{k-1}^{(1)} \left( \frac{s}{U} \right) + 1_{\{s\geq U\}} (1 - U) L_{k-1}^{(2)} \left( \frac{s - U}{1 - U} \right), \end{cases} \]

and

\[ K_1(s(1 - s))^{\beta/2} = \lim_{k \to \infty} \kappa 2^k E \left[ \prod_{\ell=1}^k X_\ell^\beta \right] E[L_k^\beta(s)], \]

implying

\[ \lim_{k \to \infty} \frac{E[L_k^\beta(s)]}{E[L_k^\beta(U)]} = \frac{K_1}{\kappa} (s(1 - s))^{\beta/2}. \]

Once again, observe that

\[ L_k(U) \overset{d}{=} \prod_{\ell=1}^k \sqrt{X_\ell} \]