Electron. J. Probab. **17** (2012), no. 42, 1–17.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2069

Distribution of the supremum location of stationary processes*

Gennady Samorodnitsky† Yi Shen‡

Abstract

The location of the unique supremum of a stationary process on an interval does not need to be uniformly distributed over that interval. We describe all possible distributions of the supremum location for a broad class of such stationary processes. We show that, in the strongly mixing case, this distribution does tend to the uniform in a certain sense as the length of the interval increases to infinity.

Keywords: stationary process ; global supremum location ; bounded variation ; strong mixing.
AMS MSC 2010: 60G10 ; 60G17.
Submitted to EJP on October 14, 2011, final version accepted on May 23, 2012.

1 Introduction

Let $\mathbf{X} = (X(t), t \in \mathbb{R})$ be a sample continuous stationary process. Even if, on an event of probability 1, the supremum of the process over a compact interval $[0, T]$ is attained at a unique point, this point does not have to be uniformly distributed over that interval, as is known since [5]. However, its distribution still has to be absolutely continuous in the interior of the interval, and the density has to satisfy very specific general constraints, as was shown in a recent paper [7].

In this paper we give a complete description of the family of possible densities of the supremum location for a large class of sample continuous stationary processes. The necessary conditions on these densities follow by combining certain general results cited above, and for every function satisfying these necessary conditions we construct a stationary process of the required type for which this function is the density of the supremum location. This is done in Section 3, which is preceded by Section 2 in which we describe the class of stationary processes we are considering and quote the results from [7] we need in the present paper. Next, we show that for a large class of stationary processes, under a certain strong mixing assumption, the distribution of the supremum location does converge to the uniformity for very long intervals, and it does it in a strong sense. This is shown in Section 4.

*Supported by grants ARO W911NF-07-1-0078, NSF DMS-1005903, NSA H98230-11-1-0154.
†Cornell University, USA. E-mail: gs18@cornell.edu
‡Cornell University, USA. E-mail: ys437@cornell.edu

2 Preliminaries

For most of this paper $\mathbf{X} = (X(t),\, t \in \mathbb{R})$ is a stationary process with continuous sample paths, defined on a probability space $\big(\Omega, \mathcal{F}, P\big)$, but in Section 4 we will allow upper semi-continuous sample paths. In most of the paper (but not in Section 4) we will also impose two assumptions on the process, which we now state.

For $T > 0$ we denote by $X_*(T) = \sup_{0\le t\le T} X(t)$, the largest value of the process in the interval $[0, T]$.

Assumption U$_T$:

$$P\Big(X(t_i) = X_*(T),\, i = 1, 2, \text{ for two different } t_1, t_2 \in [0, T]\Big) = 0.$$

Many processes satisfy Assumption U_T. In particular, a beautiful proof in [3] shows that any continuous Gaussian process, such that $X(s) \neq X(t)$ a.s. for any two points $s \neq t$, satisfies this assumption.

The second assumption on a stationary process deals with the fluctuations of its sample paths.

Assumption L:

$$K := \lim_{\varepsilon\downarrow 0} \frac{P\big(\mathbf{X} \text{ has a local maximum in } (0, \varepsilon)\big)}{\varepsilon} < \infty\,,$$

with the limit easily shown to exist. Under Assumption L the process $\mathbf{X}$ has sample paths of locally bounded variation; see Lemma 2.2 in [7].

For a compact interval $[a, b]$, we will denote by

$$\tau_{\mathbf{X},[a,b]} = \inf\big\{t \in [a, b] :\; X(t) = \max_{a\le s\le b} X(s)\big\}$$

the leftmost location of the supremum in the interval; it is a well defined random variable. If the supremum is unique, the adjective "leftmost" is, clearly, redundant. For $a = 0$, we will abbreviate $\tau_{\mathbf{X},[0,b]}$ to $\tau_{\mathbf{X},b}$, and use the same abbreviation in similar situations in the sequel.

We denote by $F_{\mathbf{X},[a,b]}$ the law of $\tau_{\mathbf{X},[a,b]}$; it is a probability measure on the interval $[a, b]$. It was proved in [7] that for any $T > 0$ the probability measure $F_{\mathbf{X},T}$ is absolutely continuous in the interior of the interval $[0, T]$, and density can be chosen to be right continuous and have left limits; we call this version of the density $f_{\mathbf{X},[a,b]}$. This version of the density satisfies a universal upper bound

$$f_{\mathbf{X},T}(t) \le \max\left(\frac{1}{t}, \frac{1}{T-t}\right),\; 0 < t < T\,. \tag{2.1}$$

We will also use the following result from the above reference.

Lemma 2.1. *Let $0 \le \Delta < T$. Then for every $0 \le \delta \le \Delta$, $f_{\mathbf{X},T-\Delta}(t) \ge f_{\mathbf{X},T}(t+\delta)$ almost everywhere in $(0, T-\Delta)$. Furthermore, for every such δ and every $\varepsilon_1, \varepsilon_2 \ge 0$, such that $\varepsilon_1 + \varepsilon_2 < T - \Delta$,*

$$\int_{\varepsilon_1}^{T-\Delta-\varepsilon_2} \big(f_{\mathbf{X},T-\Delta}(t) - f_{\mathbf{X},T}(t+\delta)\big)\, dt \tag{2.2}$$

$$\le \int_{\varepsilon_1}^{\varepsilon_1+\delta} f_{\mathbf{X},T}(t)\, dt + \int_{T-\Delta-\varepsilon_2+\delta}^{T-\varepsilon_2} f_{\mathbf{X},T}(t)\, dt\,.$$

3 Processes satisfying Assumption L

In this section we prove our main theorem, giving a full description of possible càdlàg densities $f_{\mathbf{X},T}$ for continuous stationary processes satisfying Assumption U_T and Assumption L.

For a function f of a real argument whose domain contains an interval (t_1, t_2), its total variation over the interval is defined by

$$TV_{(t_1,t_2)}(f) := \sup \sum_{i=1}^{n-1} \left| f(s_{i+1}) - f(s_i) \right|,$$

where the supremum is taken over all choices of $t_1 < s_1 < \ldots < s_n < t_2$.

Theorem 3.1. *Let $\mathbf{X} = (X(t),\, t \in \mathbb{R})$ be a stationary sample continuous process, satisfying Assumption U_T and Assumption L. Then the restriction of the law $F_{\mathbf{X},T}$ of the unique location of the supremum of the process in $[0,T]$ to the interior $(0,T)$ of the interval is absolutely continuous. The density $f_{\mathbf{X},T}$ has a càdlàg version with the following properties:*

(a) The density has a bounded variation on $(0,T)$, hence the limits

$$f_{\mathbf{X},T}(0+) = \lim_{t\to 0} f_{\mathbf{X},T}(t) \text{ and } f_{\mathbf{X},T}(T-) = \lim_{t\to T} f_{\mathbf{X},T}(t)$$

exist and are finite. Furthermore,

$$TV_{(0,T)}(f_{\mathbf{X},T}) \le f_{\mathbf{X},T}(0+) + f_{\mathbf{X},T}(T-)\,. \tag{3.1}$$

(b) The density is bounded away from zero. That is,

$$\inf_{0<t<T} f_{\mathbf{X},T}(t) > 0\,. \tag{3.2}$$

(c) Either $f_{\mathbf{X},T}(t) = 1/T$ for all $0 < t < T$, or $\int_0^T f_{\mathbf{X},T}(t)\,dt < 1$.

Moreover, if f is a nonnegative càdlàg function satisfying (a)-(c) above, then there is a stationary sample continuous process $\mathbf{X}$, satisfying Assumption U_T and Assumption L, such that f is the density in the interior $(0,T)$ of the unique location of the supremum of the process in $[0,T]$.

Proof. The existence of a càdlàg density with properties (a)-(c) in the statement of the theorem is an immediate consequence of the statements of Theorems 3.1 and 3.3 in [7]. We proceed to show the converse part of the theorem. If $f_{\mathbf{X},T}(t) = 1/T$ for all $0 < t < T$, then a required example is provided by a single wave periodic stationary Gaussian process with period T, so we need only to consider the second possibility in property (c). We start with the case where the candidate density f is a piecewise constant function of a special form.

We call a finite collection (u_i, v_i), $i = 1, \ldots, m$ of nonempty open subintervals of $(0,T)$ *a proper collection of blocks* if for any $i, j = 1, \ldots, m$ there are only 3 possibilities: either $(u_i, v_i) \subseteq (u_j, v_j)$, or $(u_j, v_j) \subseteq (u_i, v_i)$, or $[u_i, v_i] \cap [u_j, v_j] = \emptyset$. If $u_i = 0$, $v_i = T$, we call (u_i, v_i) a base block. If $u_i = 0$, $v_i < T$, we call (u_i, v_i) a left block. If $u_i > 0$, $v_i = T$, we call (u_i, v_i) a right block. If $u_i > 0$, $v_i < T$, we call (u_i, v_i) a central block. We start with constructing a stationary process as required in the theorem when the candidate density f satisfies requirements (a)-(c) of the theorem and has the form

$$f(t) = \frac{1}{HT} \sum_{i=1}^{m} \mathbf{1}_{[u_i, v_i)}(t),\ 0 < t < T \tag{3.3}$$

for some proper collection of blocks, with the obvious convention at the endpoints 0 and T, for some $H > 1$. Observe that for functions of the type (3.3), requirement (b) of the theorem is equivalent to requiring that there is at least one base block, and requirement (a) is equivalent to requiring that the number of the central blocks does not exceed the number of the base blocks. Finally, (the second case of) property (c) is equivalent to requiring that

$$d = \frac{1}{m}\left(HT - \sum_{i=1}^{m}(v_i - u_i)\right) > 0\,. \tag{3.4}$$

We will construct a stationary process by a uniform shift of a periodic deterministic function over its period. Now, however, the period will be equal to $HT > T$. We start, therefore, by defining a deterministic continuous function $(x(t),\, 0 \le t \le HT)$ with $x(0) = x(HT)$, which we then extend by periodicity to the entire $\mathbb{R}$. Let $B \ge 1$ be the number of the base blocks in the collection. We partition the entire collection of blocks into B subcollection which we call *components* by assigning each base block to one component, assigning to each component at most one central block, and assigning the left and right blocks to components in an arbitrary way. For $j = 1, \dots, B$ we denote by

$$\begin{aligned} L_j = d\big(&\text{the number of blocks in the } j\text{th component}\big) \\ &+ \text{the total length of the blocks in the } j\text{th component}\,. \end{aligned} \tag{3.5}$$

We set $x(0) = 2$. Using the blocks of the first component we will define the function x on the interval $(0, L_1]$ in such a way that $x(L_1) = 2$. Next, using the blocks of the second component we will define the function x on the interval $(L_1, L_1 + L_2]$ in such a way that $x(L_1 + L_2) = 2$, etc. Since

$$\sum_{j=1}^{B} L_j = dm + \sum_{i=1}^{m}(v_i - u_i) = HT\,,$$

this construction will terminate with a function x constructed on the entire interval $[0, HT]$ with $x(HT) = 2 = x(0)$, as desired.

We proceed, therefore, with defining the function x on an interval of length L_j using the blocks of the jth component. For notational simplicity we will take $j = 1$ and define x on the interval $[0, L_1]$ using the blocks of the first component. The construction is slightly different depending on whether or not the component has a central block, whether or not it has any left blocks, and whether or not it has any right blocks. If the component has $l \ge 1$ left blocks, we will denote them by $(0, v_j)$, $j = 1, \dots, l$. If the component has $r \ge 1$ right blocks, we will denote them by (u_j, T), $j = 1, \dots, r$. If the component has a central block, we will denote it by (u, v). We will construct the function x by defining it first on a finite number of special points and then filling in the gaps in a piecewise linear manner.

Suppose first that the component has a central block, some left blocks and some right blocks. In this case we proceed as follows.

Step 1 Recall that $x(0) = 2$ and set

$$x\left(jd + \sum_{i=1}^{j-1} v_i\right) = x\left(jd + \sum_{i=1}^{j} v_i\right) = 2 - 2^{j-l},\ j = 1, \dots, l\,.$$

Note that the last point obtained in this step is $x\big(ld + \sum_{i=1}^{l} v_i\big) = 1$.

Step 2 Set

$$x\left((l+1)d + \sum_{i=1}^{l} v_i\right) = x\left((l+1)d + \sum_{i=1}^{l} v_i + v\right)$$

$$= x\left((l+1)d + \sum_{i=1}^{l} v_i + v + T - u\right) = \frac{1}{2}\,.$$

Step 3 Set

$$x\left((l+j+1)d + \sum_{i=1}^{l} v_i + v + T - u + \sum_{i=1}^{j-1}(T-u_j)\right)$$

$$= x\left((l+j+1)d + \sum_{i=1}^{l} v_i + v + T - u + \sum_{i=1}^{j}(T-u_j)\right)$$

$$= 2 - 2^{-(j-1)},\ j = 1, \ldots, r\,.$$

Note that the last point obtained in this step is

$$x\left((l+r+1)d + \sum_{i=1}^{l} v_i + v + T - u + \sum_{i=1}^{r}(T-u_j)\right) = 2 - 2^{-(r-1)}\,.$$

Step 4 We add just one more point at distance d from the last point of the previous step by setting

$$x\left((l+r+2)d + \sum_{i=1}^{l} v_i + v + T - u + \sum_{i=1}^{r}(T-u_j)\right) = 2\,.$$

Note that this point coincides with L_1 as defined in (3.5).

If the component has no left blocks, then Step 1 above is skipped, and Step 2 becomes the initial step with

$$x(d) = x(d+v) = x(d+v+T-u) = \frac{1}{2}\,.$$

If the component has no right blocks, then Step 3 above is skipped, and at Step 4 we add the distance d to the final point of Step 2, that is we set

$$x\left((l+2)d + \sum_{i=1}^{l} v_i + v + T - u\right) = 2\,.$$

If the component has no central block, then Step 2 is skipped, but we do add the distance T to the last point of Step 1. That is, the first point obtained at Step 3 becomes

$$x\left((l+1)d + \sum_{i=1}^{l} v_i + T\right) = 1\,,$$

if there are any left blocks, with the obvious change if $l = 0$. Finally, if there is neither central block, nor any right blocks, then both Step 2 and Step 3 are skipped, and Step 4 just adds $d + T$ to the last point of Step 1, i.e. it becomes

$$x\left((l+1)d + \sum_{i=1}^{l} v_i + T\right) = 2\,,$$

once again with the obvious change if $l = 0$. It is easy to check that in any case Step 4 sets $x(L_1) = 2$, with L_1 as defined in (3.5). In particular, $L_1 > T$.

Finally, we specify the piecewise linear rule by which we complete the construction of the function x on the interval $[0, L_1]$. The function has been defined on a finite set of points and we proceed from left to right, starting with $x(0) = 2$, to fill the gap between one point in the finite set and the adjacent point from the right, until we reach $x(L_1) = 2$. By the construction, there are pairs of adjacent points in which the values of x coincide, and pairs of adjacent points in which the values of x are different. In most cases only adjacent points at the distance d have equal values of x, but if, e.g. a central block is missing, then at a pair of adjacent points at a distance T, or $d + T$, the values of x coincide as well.

In any case, if the values of x at two adjacent points are different, we define the values of x between these two points by linear interpolation. If the values of x at two adjacent points, say, a and b with $a < b$, are equal to, say, y we define the function x between these two points by

$$x(t) = \max\big(y - (t-a)/d,\, y - (b-t)/d\big)$$

provided the value at the midpoint, $y - (b-a)/2d \geq -1$. If this lower bound fails, we define the values of x between the points $a + dy$ and $b - dy$ by

$$x(t) = \max\big(-\tau(t - (a+dy)),\, -\tau((b-dy) - t)\big)\,,$$

for an arbitrary $\tau > 0$ such that both $\tau \leq 1/d$ and the value at the midpoint, $-\tau\big((b-a)/2 - dy\big) \geq -1$. The reason for this slightly cumbersome definition is the need to ensure that x is nowhere constant, while keeping the lower bound of x and its Lipschitz constant under control. We note, at this point, that, since in all cases $b - a \leq T + d$, we can choose, for a fixed T, the value of τ so that $\tau \geq \tau_d > 0$, where the constant τ_d stays bounded away from zero for d in a compact interval.

Now that we have defined a periodic function $(x(t),\, t \in \mathbb{R})$ with period HT, we define a stationary process $\mathbf{X}$ by $X(t) = x(t - U)$, $t \in \mathbb{R}$, where U is uniformly distributed between 0 and HT. The process is, clearly, sample continuous and satisfies Assumption L. We observe, further, that, if the supremum in the interval $[0, T]$ is achieved in the interior of the interval, then it is achieved at a local maximum of the function x. If the value at the local maximum is equal to 2, then it is due to an endpoint of a component, and, since the contribution of any component has length exceeding T, this supremum is unique. If the value at the local maximum is smaller than 2, then that local maximum is separated from the nearest local maximum with the same value of x by at least the distance induced by Step 2, which T. Consequently, in this case the supremum over $[0, T]$ is unique as well. Similarly, if the supremum is achieved at one of the endpoints of the interval, it has to be unique as well, on a set of probability 1. Therefore, the process $\mathbf{X}$ satisfies Assumption U_T.

Example We interrupt the exposition for a moment to demonstrate a simple special case of the construction of the process $\mathbf{X}$ to help the reader to visualize the procedure. Consider a candidate density function

$$f(t) = \begin{cases} \frac{2}{HT} & \text{if } t \in (0, v_1) \cup [u, v), \\ \frac{1}{HT} & \text{if } t \in [v_1, u) \cup [v, T) \end{cases}$$

for $0 < v_1 < u < v < T$, with $H > 1 + \frac{v_1 + v - u}{T}$. This corresponds to a proper collection of three blocks: a base block $(0, T)$, a central block (u, v), and a left block $(0, v_1)$. Hence the total number of blocks $m = 3$, and $d = \frac{1}{3}(HT - T - v_1 - (v - u)) > 0$. Since there is only one base block, we use one component, of the length $L_1 = HT$.

The construction of the deterministic function $x(t)$ on $[0, HT]$ is as follows. The starting point is $x(0) = 2$. Then step 1, dealing with the left block $(0, v_1)$, assigns value

$2 - 2^{1-1} = 1$ to points d and $d + v_1$. Step 2 continues to set

$$x(2d + v_1) = x(2d + v_1 + v) = x(2d + v_1 + v + T - u) = \frac{1}{2}.$$

Step 3 is skipped since there is no right block. Finally, the end point of this component, $x(3d + v_1 + v + T - u) = 2$ is added in step 4. Since $3d + v_1 + v + T - u = HT$, this is the end of the cycle.

To demonstrate the the piecewise linear interpolation rule between these special points, we choose specific values $T = 6, H = 2, v_1 = 1, u = 3$ and $v = 5$. This implies $d = 1$. Firstly, between the pairs of points with the t coordinates 0 and $d = 1$, $d + v_1 = 2$ and $2d + v_1 = 3$, $2d + v_1 + v + T - u = 11$ and $HT = 12$ we use linear interpolation. Consider the segment between the points $d = 1$ and $d + v_1 = 2$, at which x has the common value $y = 1$. The general rule checks the value of the interpolation at the midpoint of the segment, which is $1 - \frac{v_1}{2d} = 1/2$. It is greater than -1, so no modification is necessary. Same procedure applies to the segments between the points $2d + v_1 = 3$ and $2d + v_1 + v = 8$, and the points $2d + v_1 + v = 8$ and $2d + v_1 + v + T - u = 11$. Only on the interval $(3, 8)$ the interpolation procedure has to be modified. We set $\tau = 1/2$ (so that the value of the lowest point is exactly -1) and obtain

$$x(t) = \begin{cases} 3.5 - t & \text{if } 3 \le t \le 3.5 \\ (3.5 - t)/2 & \text{if } 3.5 \le t \le 5.5 \\ (7.5 - t)/2 & \text{if } 5.5 \le t \le 7.5 \\ 7.5 - t & \text{if } 7.5 \le t \le 8 \end{cases}.$$

The figure below shows the density f and the function $x(t)$ within one cycle.

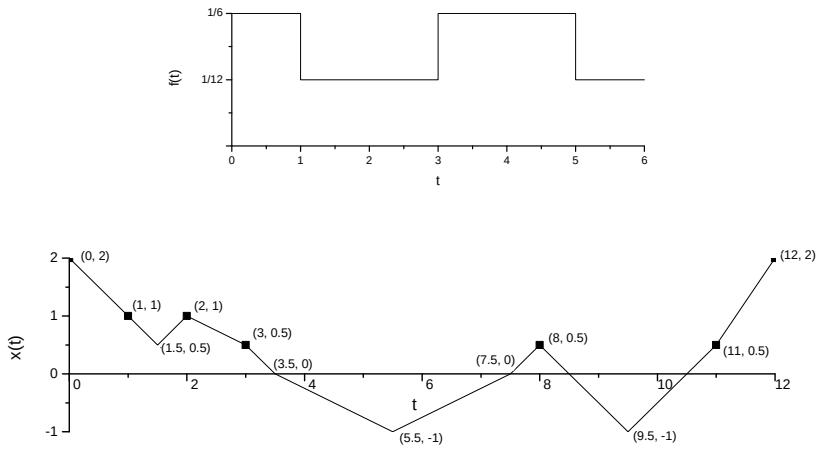

Fig 1. The functions f and x in the special case.

Finally, we extend the function $x(t)$ periodically with period HT to the whole real line. The process $\mathbf{X}$ is then defined by $\mathbf{X}(t) = x(t - U)$, where U is uniformly distributed between 0 and HT.

We now return to the general case considered in the theorem. We first show that for the process $\mathbf{X}$ constructed above, the density $f_{\mathbf{X},T}$ coincides with the function f given in (3.3), with which the construction was performed. According to the above analysis, we need to account for the contribution of each local maximum of the function x over its period to the density $f_{\mathbf{X},T}$. The local maxima may appear in Step 1 of the construction, and then they are due to left blocks. They may apear in Step 3 of the construction, and then they are due to right blocks. They may appear Step 2 of the construction,

and then they are due to central blocks. Finally, the points where x has value 2 are always local maxima. We will see that they are due to base blocks. We start with the latter local maxima. Clearly, each such local maximum is, by periodicity, equal to one of the B values, $\sum_{j=1}^{i} L_j - HT$, $i = 1, \ldots, B$. The ith of these points becomes the global maximum of $\mathbf{X}$ over $[0, T]$ if and only if

$$U \in \left(HT - \sum_{j=1}^{i} L_j,\ (H+1)T - \sum_{j=1}^{i} L_j \right),$$

and the global maximum is then located at the point $\sum_{j=1}^{i} L_j - HT + U$. Therefore, the contribution of each such local maximum to the density is $1/HT$ at each $0 < t < T$, and overall the points where x has value 2 contribute to $f_{\mathbf{X},T}$

$$f_{\text{base}}(t) = \frac{B}{HT},\ 0 < t < T\,. \tag{3.6}$$

Next, we consider the contribution to $f_{\mathbf{X},T}$ of the local maxima due to left blocks. For simplicity of notation we consider only the left blocks in the first component. Then the local maximum due to the jth left block is at the point $jd + \sum_{i=1}^{j} v_i$. As before, we need to check over what interval of the values of U this local maximum becomes the global maximum of $\mathbf{X}$ over $[0, T]$. The relevant values of U must be such that the time interval $\left(jd + \sum_{i=1}^{j-1} v_i,\ jd + \sum_{i=1}^{j} v_i\right)$ is shifted to cover the origin, and this corresponds to an interval of length v_j of the values of U. The shifted local maximum itself will then be located within the interval $(0, v_j)$, which contributes $1/HT$ at each $0 < t < v_j$. Overall, the local maxima due to left blocks contribute to $f_{\mathbf{X},T}$

$$f_{\text{left}}(t) = \frac{1}{HT} \sum_{\text{left blocks}} \mathbf{1}_{(0,v_i)}(t),\ 0 < t < T\,. \tag{3.7}$$

Similarly, the local maxima due to right blocks contribute to $f_{\mathbf{X},T}$

$$f_{\text{right}}(t) = \frac{1}{HT} \sum_{\text{right blocks}} \mathbf{1}_{(u_i,T)}(t),\ 0 < t < T\,. \tag{3.8}$$

Finally, we consider the central blocks. If the first component has a central block, then the local maximum due to the central block is at the point $(l+1)d + \sum_{i=1}^{l} v_i + v$. Any value of U that makes this local maximum the global maximum over $[0, T]$ must be such that the time interval $\left((l+1)d + \sum_{i=1}^{l} v_i,\ (l+1)d + \sum_{i=1}^{l} v_i + v\right)$ is shifted to cover the origin. Furthermore, that value of U must also be such that the time interval $\left((l+1)d + \sum_{i=1}^{l} v_i + v,\ (l+1)d + \sum_{i=1}^{l} v_i + v + T - u\right)$ is shifted to cover the right endpoint T. If we think of shifting the origin instead of shifting x, the origin will have to be located in the interval $\left((l+1)d + \sum_{i=1}^{l} v_i,\ (l+1)d + \sum_{i=1}^{l} v_i + v - u\right)$. This corresponds to a set of values of U of measure $v - u$, and the shifted local maximum will then be located within the interval (u, v), which contributes $1/HT$ at each $u < t < v$ to the density. Overall, the local maxima due to central blocks contribute to $f_{\mathbf{X},T}$

$$f_{\text{central}}(t) = \frac{1}{HT} \sum_{\text{central blocks}} \mathbf{1}_{(u,v)}(t),\ 0 < t < T\,. \tag{3.9}$$

Since

$$f_{\mathbf{X},T}(t) = f_{\text{base}}(t) + f_{\text{left}}(t) + f_{\text{right}}(t) + f_{\text{central}}(t),\ 0 < t < T\,,$$

we conclude by (3.6) - (3.9) that $f_{\mathbf{X},T}$ indeed coincides with the function f given in (3.3). Therefore, we have proved the converse part of the theorem in the case when the candidate density f is of the form (3.3).

We now prove the converse part of the theorem for a general f with properties (a)-(c) in the statement of the theorem. Recall that we need only to treat the second possibility in property (c). In order to construct a stationary process $\mathbf{X}$ for which $f_{\mathbf{X},T} = f$, we will approximate the candidate density f by functions of the form (3.3). Since we will need to deal with convergence of a sequence of continuous stationary processes we have just constructed in the case when the candidate density is of the form (3.3), we record, at this point, several properties of the stationary periodic process $X(t) = x(t-U)$, $t \in \mathbb{R}$ constructed above.

Property 1 *The process* $\mathbf{X}$ *is uniformly bounded:* $-1 \le X(t) \le 2$ *for all* $t \in \mathbb{R}$.

Property 2 *The process* $\mathbf{X}$ *is Lipschitz continuous, and its Lipschitz constant does not exceed* $3/2d$.

Property 3 *The process* $\mathbf{X}$ *is differentiable except at countably many points, at which* $\mathbf{X}$ *has left and right derivatives. On the set* $D_0 = \{t : X(t) > 0\}$ *the derivatives satisfy*

$$|X'(t)| \ge \frac{1}{2^N d}$$

(where the bound applies to both left and right derivatives if t *is not a differentiability point). Here* N *is the bigger of the largest number of left blocks any component has, and the largest number of the right blocks any component has. Similarly, on the set* $D_1 = \{t : X(t) \le 0\}$ *the derivatives satisfy*

$$|X'(t)| \ge \tau_d\,,$$

where $\tau_d > 0$ *stays bounded away from zero for* d *in a compact interval.*

Property 4 *The distance between any two local maxima of* $\mathbf{X}$ *cannot be smaller than* d. *At its local maxima,* $\mathbf{X}$ *takes values in a finite set of at most* $N+3$ *elements. Moreover, the absolute difference in the values of the process* $\mathbf{X}$ *in two local maxima in the interval* $(0,T)$ *is at least* 2^{-N}, *where* N *is as above.*

All these properties follow from the corresponding properties of the function x by considering the possible configuration of the blocks in a component.

We will now construct a sequence of approximations to a candidate density f as above. Let $n = 1, 2, \ldots$. It follows from the general properties of càdlàg functions (see e.g. [1]) that there is a finite partition $0 = t_0 < t_1 < \ldots < t_k = T$ of the interval $[0,T]$ such that

$$|f(s) - f(t)| \le \frac{1}{nT} \text{ for all } t_i \le s, t < t_{i+1},\ i = 0, \ldots, k-1. \tag{3.10}$$

We define a piecewise constant function $\tilde{f}_n$ on $(0,T)$ by setting, for each $i = 1, \ldots, k$, the value of $\tilde{f}_n$ for $t_{i-1} \le t < t_i$ to be

$$\tilde{f}_n(t) = \frac{1}{knT} \max\left\{ j = 0, 1, \ldots : f(s) \ge \frac{j}{knT} \text{ for all } t_{i-1} \le s < t_i \right\}.$$

By definition and (3.10) we see that

$$f(t) - \frac{2}{nT} \le \tilde{f}_n(t) \le f(t),\ 0 < t < T\,. \tag{3.11}$$

Next, we notice that for every $i = 1, \ldots, k-1$ there are points $s_i \in (t_{i-1}, t_i)$ and $s_{i+1} \in (t_i, t_{i+1})$ such that

$$\left|f(s_i) - f(s_{i+1})\right| \ge \left|\tilde{f}_n(t_i-) - \tilde{f}_n(t_i)\right| - \frac{1}{knT}\,.$$

Therefore,

$$TV_{(0,T)}(\tilde{f}_n) \le TV_{(0,T)}(f) + \frac{1}{nT}\,. \tag{3.12}$$

We now define

$$f_n(t) = \tilde{f}_n(t) + \frac{1}{nT}\ \ 0 < t < T\,.$$

Clearly, the function f_n is càdlàg, has bounded variation on $(0,T)$ and is bounded away from zero. By (3.12), f_n also satisfies (3.1) since f does. Finally, since $\int_0^T f_{\mathbf{X},T}(t)\,dt < 1$, we see by (3.11) that, for all n large enough, $\int_0^T f_{\mathbf{X},T}(t)\,dt < 1$ as well. Therefore, for such n the function f_n has properties (a)-(c) in the statement of the theorem, and in the sequel we will only consider n large as above. We finally notice that f_n takes finitely many different values, all of which are in the set $\{j/knT,\ j = 1, 2, \ldots\}$. Therefore, f_n can be written in the form (3.3), with $H = kn$. Indeed, the blocks can be built by combining into a block all neighboring intervals where the value of f_n is the smallest, subtracting $1/knT$ from the value of f_n in the constructed block and iterating the procedure.

We have already proved that for any function of the type (3.3) there is a stationary process required in the statement of the theorem. Recall that a construction of this stationary process depends on assignment of blocks in a proper collection to components, and we would like to make sure that no component has "too many" left or right blocks. To achieve this, we need to distribute the left and right blocks as evenly as possible between the components. Two observations are useful here. First of all, it follows from the definition of f_n and (3.3) that

$$\frac{1}{k_n nT}(L_n + B_n) = f_n(0+) \le f(0+) + \frac{1}{k_n nT} \le f(0+) + 1$$

for n large enough (we are writing k_n instead of k to emphasize the dependence of k on n), where L_n and B_n are the numbers of the the left and base blocks in the nth collection. On the other hand, similar considerations tell us that

$$\frac{1}{k_n nT}B_n = \inf_{0<t<T} f_n(t) \ge \inf_{0<t<T} f(t) - \frac{2}{nT} \ge \frac{1}{2}\inf_{0<t<T} f(t)\,,$$

once again for n large enough, where we have used property (b) of f. Therefore, for such n,

$$\frac{L_n}{B_n} \le 2\frac{f(0+)+1}{\inf_{0<t<T} f(t)}\,, \tag{3.13}$$

and the right hand side is a finite quantity depending on f, but not on n. Performing a similar analysis for the right blocks, and recalling that we are distributing the left and right blocks as evenly as possible between the components, we see that there is a number $\Delta_f \in (0,\infty)$ such that for all n large enough, no component in the nth collection has more than Δ_f left blocks or Δ_f right blocks.

We will also need bounds on the important parameter $d = d_n$ appearing in the construction of a stationary process corresponding to functions of the type (3.3); these bounds do not depend on a particular way we assigns blocks to different components. Recall that

$$d_n = \frac{k_n nT}{m_n}\left(1 - \int_0^T f_n(t)\,dt\right)\,, \tag{3.14}$$

where $m_n = B_n + L_n + R_n + C_n$ (in the obvious notation) is the total number of blocks in the nth collection. Since

$$\frac{1}{k_n nT}\big(B_n + \max(L_n, R_n, C_n)\big) = \sup_{0<t<T} f_n(t),\quad \frac{1}{k_n nT}B_n = \inf_{0<t<T} f_n(t)\,,$$

we see that

$$\inf_{0<t<T} f_n(t) \le \frac{1}{k_n n T} m_n \le 3 \sup_{0<t<T} f_n(t)\,. \tag{3.15}$$

We also know by the uniform convergence that $\int_0^T f_n \to \int_0^T f$. Therefore, by (3.14) and (3.15) we obtain that, for all n large enough,

$$\frac{1-\int_0^T f(t)\,dt}{4\sup_{0<t<T} f(t)} \le d_n \le \frac{2}{\inf_{0<t<T} f(t)}\,. \tag{3.16}$$

An immediate conclusion is the following fact. By construction, the distribution of $X_n(0)$ is absolutely continuous; let g_n denote the right continuous version of its density. Since $\mathbf{X}_n$ is obtained by uniform shifting of a piecewise linear periodic function with period $H_n T$, the value of the density $g_n(v)$ at each point v times the length of the period does not exceed the total number of the linear pieces in a period divided by the smallest absolute slope of any linear piece. The former does not exceed $2m_n$, and by **Property 3** and the above, the latter cannot be smaller than

$$\min\left(\frac{1}{2^{\Delta_f} d_n},\, \tau_{d_n}\right)\,.$$

Since, by (3.16), d_n is uniformly bounded from above, we conclude, for some finite positive constant $c = c(f)$, $g_n(v) \le c(f) m_n/H_n$. Further, by the definition of d_n,

$$m_n d_n = H_n T P\big(\tau_{\mathbf{X}_n,T} \in \{0,T\}\big) \le H_n T\,.$$

Once again, since by (3.16), d_n is uniformly bounded from below, we conclude that

$$g_n(v) \text{ is uniformly bounded in } v \text{ and } n. \tag{3.17}$$

Let $\mathbf{X}_n$ be the stationary process corresponding to f_n constructed above. We view $\mathbf{X}_n$ as a random element of the space $C(\mathbb{R})$ of continuous functions on $\mathbb{R}$ which we endow with the metric

$$\rho(\mathbf{x},\mathbf{y}) = \sum_{m=1}^{\infty} 2^{-m}\big(\sup_{|t|\le m} |x(t)-y(t)|\big)\,.$$

Let μ_n be the law of $\mathbf{X}_n$ on $C(\mathbb{R})$, $n = 1,2,\dots$ (but large enough, as needed). By **Property 1** and **Property 2** of the processes $\mathbf{X}_n$ and the lower bound in (3.16), these processes are uniformly bounded and equicontinuous. Therefore, by Theorem 7.3 in [1], for every fixed $m = 1,2,\dots$ the restrictions of the measures μ_n to the interval $[-m,m]$ form a tight family of probability measures. Let $n_{1j} \to \infty$ be a sequence positive integers such that the restrictions of $\mu_{n_{1j}}$ to $[-1,1]$ converge weakly to a probability measure ν_1 on $C([-1,1])$. Inductively define for $m = 2,3,\dots$ $n_{mj} \to \infty$ to be a subsequence of the sequence $n_{m-1,j} \to \infty$ such that the restrictions of $\mu_{n_{mj}}$ to $[-m,m]$ converge weakly to a probability measure ν_m on $C([-m,m])$. Then the "diagonal" sequence of measures $\big(\mu_{n_{jj}},\, j = 1,2,\dots\big)$ is such that the restrictions of these measures to each interval $[-m,m]$ converge weakly to ν_m on $C([-m,m])$. By the Kolmogorov existence theorem, there is a (cylindrical) probability measure ν on functions on $\mathbb{R}$ whose restrictions to each interval $[-m,m]$ coincide with ν_m (considered now as a cylindrical measure). Since each probability measure ν_m is supported by $C([-m,m])$, the measure ν itself is supported by functions in $C(\mathbb{R})$. By construction, the measure ν is shift invariant. If $\mathbf{X}$ is the canonical stochastic process defined on $\big(C(\mathbb{R}),\nu\big)$, then $\mathbf{X}$ is a sample continuous stationary process. In the remainder of the proof we will show that $\mathbf{X}$ satisfies Assumption L and Assumption U_T, and that $f_{\mathbf{X},T} = f$.

We start with proving that Assumption L holds for $\mathbf{X}$. It is, clearly, enough to prove that, on a set of probability 1,

$$\text{any two local maxima of } \mathbf{X} \text{ are at least } \theta := \frac{1 - \int_0^T f(t)\,dt}{5 \sup_{0<t<T} f(t)} \text{ apart.} \tag{3.18}$$

Suppose that (3.18) fails. Then there is m sucht that, on an event of positive probability, two local maxima of $\mathbf{X}$ closer than θ exist in the time interval $[-m, m]$. Recall that a subsequence of the sequence of the (laws of) $\mathbf{X}_n$ converges weakly in the uniform topology on $C([-m, m])$ to the (law of) $\mathbf{X}$. For notational simplicity we will identify that subsequence with the entire sequence $(\mathbf{X}_n)$. By the Skorohod representation theorem (Theorem 6.7 in [1]), we may define the processes $(\mathbf{X}_n)$ on some probability space so that $\mathbf{X}_n \to \mathbf{X}$ a.s. in $C([-m, m])$. Fix ω for which this convergence holds, and for which $\mathbf{X}$ has two local maxima closer than θ exist in the time interval $[-m, m]$. It straightforward to check that the uniform convergence and **Property 3** above imply that for all n large enough, the processes $\mathbf{X}_n$ will have two local maxima closer than $5\theta/4$. This is, of course, impossible, due to **Property 4** and (3.16). The resulting contradiction proves that $\mathbf{X}$ satisfies Assumption L.

Next, we prove that Assumption U_T holds for $\mathbf{X}$. Since the process $\mathbf{X}$ satisfies Assumption L, by Lemma 2.2 in [7], it has finitely many local maxima in the interval $(0, T)$ (in fact, by (3.18), it cannot have more than $\lceil T/\theta \rceil$ local maxima). Clearly, the values of $\mathbf{X}$ at the largest local maximum and the second largest local maximum (if any) are well defined random variables. We denote by (M_1, M_2) the largest and the second largest among $X(0)$, $X(T)$ and the values of $\mathbf{X}$ at the largest local maximum and the second largest local maximum (if any). The fact that Assumption U_T holds for $\mathbf{X}$ will follow once we prove that

$$P\big(M_1 = M_2\big) = 0\,. \tag{3.19}$$

We proceed similarly to the argument in the proof of Assumption L. We may assume that $\mathbf{X}_n \to \mathbf{X}$ a.s. in $C[0, T]$. Fix ω for which this convergence holds. The uniform convergence and **Property 3** of the processes $(\mathbf{X}_n)$, together with the uniform upper bound on d_n in (3.16), show that, for every local maximum t_ω of $\mathbf{X}$ in the interval $(0, T)$ and any $\delta > 0$, there is $n(\omega, \delta)$ such that for all $n > n(\omega, \delta)$, the process $\mathbf{X}_n$ has a local maximum in the interval $(t_\omega - \delta, t_\omega + \delta)$. This immediately implies that

$$M_1 - M_2 \geq \limsup_{n\to\infty} \big(M_1^{(n)} - M_2^{(n)}\big)$$

a.s., where the random vector $(M_1^{(n)}, M_2^{(n)})$ is defined for the process $\mathbf{X}_n$ in the same way as the random vector (M_1, M_2) is defined for the process $\mathbf{X}$, $n = 1, 2, \ldots$. In particular, for any $\varepsilon > 0$,

$$P\big(M_1 - M_2 < \varepsilon\big) \leq \limsup_{n\to\infty} P\big(M_1^{(n)} - M_2^{(n)} < \varepsilon\big)\,. \tag{3.20}$$

As a first step, notice that, by **Property 4** of the processes $(\mathbf{X}_n)$, for any $\varepsilon < \Delta_f$,

$$P\Big(M_1^{(n)} - M_2^{(n)} < \varepsilon, \tag{3.21}$$

$$\text{both } M_1^{(n)} \text{ and } M_2^{(n)} \text{ achieved at local maxima}\Big) = 0$$

for each n. Next, since by **Property 4**, at its local maxima the process $\mathbf{X}_n$ can take at most $\Delta_f + 3$ possible values, we conclude by (3.17) that for all $\varepsilon > 0$,

$$P\Big(M_1^{(n)} - M_2^{(n)} < \varepsilon, \text{ one of } M_1^{(n)} \tag{3.22}$$

$$\text{and } M_2^{(n)} \text{ is achieved at a local maximum, and one at an endpoint}\Big) \le c_f \varepsilon\,,$$

for some $c_f \in (0,\infty)$. Finally, we consider the case when both $M_1^{(n)}$ and $M_2^{(n)}$ are achieved at the endpoints of the interval $[0,T]$. In that case, it is impossible that $\mathbf{X}_n$ has a local maximum in $(0,T)$, since that would force time 0 to belong to one of the decreasing linear pieces of the process due to left blocks, and time T to belong one of the increasing linear pieces of the process due to right blocks. By construction, the distance between any two points belonging to such intervals is larger than T. That forces $X_n(t), 0 \le t \le T$ to consist of at most two linear pieces. By **Property 3** of the process $\mathbf{X}_n$, in order to achieve $|X_n(0) - X_n(T)| \le \varepsilon$, each block of the proper collection generating $\mathbf{X}_n$ contributes at most an interval of length $\varepsilon/\min(1/(2^\Delta d_n), \tau_{d_n})$ to the set of possible shifts U. Recall that there are m_n blocks in the collection. By the uniform bounds (3.16) we conclude that for all $\varepsilon > 0$,

$$P\Big(M_1^{(n)} - M_2^{(n)} < \varepsilon, \tag{3.23}$$

$$M_1^{(n)} \text{ and } M_2^{(n)} \text{ achieved at the endpoints}\Big)$$

$$\le \varepsilon \frac{m_n}{H_n T} \frac{1}{\min(1/(2^\Delta d_n), \tau_{d_n})} \le \varepsilon \frac{1}{d_n \min(1/(2^\Delta d_n), \tau_{d_n})} \le \tilde{c}_f \varepsilon\,,$$

for some $\tilde{c}_f \in (0,\infty)$.

Combining (3.20), (3.21), (3.22) and (3.23) we see that for all $\varepsilon > 0$ small enough,

$$P\big(M_1 - M_2 < \varepsilon\big) \le (c_f + \tilde{c}_f)\varepsilon\,.$$

Letting $\varepsilon \downarrow 0$ we obtain (3.19), so that the process $\mathbf{X}$ satisfies Assumption U_T.

It is now a simple manner to finish the proof of the theorem. Assume, once again, that $\mathbf{X}_n \to \mathbf{X}$ a.s. in $C[0,T]$. Fix ω for which this convergence holds, and both $\mathbf{X}$ and each $\mathbf{X}_n$ have a unique supremum in the interval $[0,T]$. It follows from the uniform convergence that $\tau_{\mathbf{X}_n,T} \to \tau_{\mathbf{X},T}$ as $n \to \infty$. Therefore, we also have that $\tau_{\mathbf{X}_n,T} \Rightarrow \tau_{\mathbf{X},T}$ (weakly). However, by construction, $f_n(t) \to f(t)$ for every $0 < t < T$. This implies that f is the density of $\tau_{\mathbf{X},T}$, and the proof of the theorem is complete. □

Remark 3.2. Recall that most of the work in the construction of the process in Theorem 3.1 was done for the second case of condition (c). Under assumptions U_T and L this is the only alternative to the uniform distribution of the location of the supremum, and it guarantees that the probability that the supremum is located at an endpoint of the interval is positive. The separation of the local maxima property (3.18) of the constructed process, for example, shows that that process satisfies assumption L. Under assumptions U_T and L, the only scenario leading to the uniform distribution is to have the global suprema of the process appear periodically with a period of T:

$$P\left(X(\tau_{\mathbf{X},[T,2T]}) = X(\tau_{\mathbf{X},T}), \tau_{\mathbf{X},[T,2T]} - \tau_{\mathbf{X},T} = T\right) = 1;$$

see [7]. It is an open problem to describe all possible scenarios leading to the uniform distribution when Assumption L no longer holds, but the dichotomy of condition (c) in Theorem 3.1 no longer needs to hold in this case, as the example of the stationary Ornstein-Uhlenbeck process considered in [7] shows.

4 Long intervals

In spite of the broad range of possibilities for the distribution of the supremum location shown in the previous section, it turns out that, when the length of an interval

becomes large, and the process satisfies a certain strong mixing assumption, uniformity of the distribution of the supremum location becomes visible at certain scales. We make this statement precise in this section.

In this section we allow a stationary process $\mathbf{X}$ to have upper semi-continuous, not necessarily continuous, sample paths. Moreover, we will not generally impose either Assumption U_T, or Assumption L. Without Assumption U_T, the supremum may not be reached at a unique point, so we will work with the leftmost supremum location defined in Section 2.

Recall that a stationary stochastic process $\mathbf{X} = (X(t), t \in \mathbb{R})$ is called strongly mixing (or α-mixing) if

$$\sup\Big\{\big|P(A\cap B) - P(A)P(B)\big| : A \in \sigma\big(X(s),\, s\le 0\big),\, B\in\sigma\big(X(s),\, s\ge t\big)\Big\}$$
$$\to 0 \text{ as } t\to\infty;$$

see e.g. [6], p. 195. Sufficient conditions on the spectral density of a stationary Gaussian process that guarantee strong mixing were established in [4].

Let $\mathbf{X}$ be an upper semi-continuous stationary process. We introduce a "tail version" of the strong mixing assumption, defined as follows.

Assumption TailSM: there is a function $\varphi : (0,\infty)\to\mathbb{R}$ such that

$$\lim_{t\to\infty} P\big(\sup_{0\le s\le t} X(s)\ge \varphi(t)\big) = 1$$

and

$$\sup\Big\{\big|P(A\cap B) - P(A)P(B)\big| : A \in \sigma\big(X(s)\mathbf{1}(X(s)\ge\varphi(t)),$$
$$s\le 0\big),\, B\in\sigma\big(X(s)\mathbf{1}(X(s)\ge\varphi(t)),\, s\ge t\big)\Big\} \to 0 \text{ as } t\to\infty.$$

It is clear that if a process is strongly mixing, then it also satisfies Assumption TailSM. The point of the latter assumption is that we are only interested in mixing properties of the part of the process "responsible" for its large values. For example, the process

$$X(t) = \begin{cases} Y(t) & \text{if } Y(t) > 1 \\ Z(t) & \text{if } Y(t)\le 1 \end{cases},\ t\in\mathbb{R}\,,$$

where $\mathbf{Y}$ is a strongly mixing process such that $P(Y(0) > 1) > 0$, and $\mathbf{Z}$ an arbitrary stationary process such that $P(Z(0) < 1) = 1$, does not have to be strongly mixing, but it clearly satisfies Assumption TailSM with $\varphi\equiv 1$.

We will impose one more assumption on the stationary processes we consider in this section. It deals with the size of the largest atom the distribution of the supremum of the process may have.

Assumption A:

$$\lim_{T\to\infty}\sup_{x\in\mathbb{R}} P\big(\sup_{t\in[0,T]} X(t) = x\big) = 0\,.$$

In Theorem 4.1 below Assumption A could be replaced by requiring Assumption U_T for all T large enough. We have chosen Assumption A instead since for many important stationary stochastic processes the supremum distribution is known to be atomless anyway; see e.g. [8] for continuous Gaussian processes and [2] for certain stable processes. The following sufficient condition for Assumption A is also elementary: suppose that the process $\mathbf{X}$ is ergodic. If for some $a\in\mathbb{R}$, $P\big(\sup_{t\in[0,1]} X(t) = x\big) = 0$ for all $x > a$ and $P(X(0) > a) > 0$, then Assumption A is satisfied.

Theorem 4.1. *Let $\mathbf{X} = (X(t),\, t \in \mathbb{R})$ be a stationary sample upper semi-continous process, satisfying Assumption TailSM and Assumption A. The density $f_{\mathbf{X},T}$ of the supremum location satisfies*

$$\lim_{T\to\infty} \sup_{\varepsilon \le t \le 1-\varepsilon} \Big| T f_{\mathbf{X},T}(tT) - 1 \Big| = 0 \tag{4.1}$$

for every $0 < \varepsilon < 1/2$. In particular, the law of $\tau_{\mathbf{X},T}/T$ converges weakly to the uniform distribution on $(0,1)$.

Proof. It is obvious that (4.1) implies weak convergence of the law of $\tau_{\mathbf{X},T}/T$ to the uniform distribution. We will, however, prove the weak convergence first, and then use it to derive (4.1).

We start with a useful claim that, while having nothing to do with any mixing by itself, will be useful for us in a subsequent application of Assumption TailSM. Let T_n, $d_n \uparrow \infty$, $d_n/T_n \to 0$ as $n \to \infty$. We claim that for any $\delta \in (0,1)$,

$$P\Big(\delta T_n - d_n \le \tau_{\mathbf{X},T_n} \le \delta T_n + d_n\Big) = 0\,. \tag{4.2}$$

To see this, simply note that by (2.1), the probability in (4.2) is bounded from above by

$$2d_n \sup_{\delta T_n - d_n \le t \le \delta T_n + d_n} f_{\mathbf{X},T_n}(t) \le 2d_n \max\left(\frac{1}{\delta T_n - d_n},\, \frac{1}{(1-\delta)T_n - d_n}\right) \to 0$$

as $n \to \infty$.

The weak convergence stated in the theorem will follow once we prove that for any rational number $r \in (0,1)$, we have $P\big(\tau_{\mathbf{X},T} \le rT\big) \to r$ as $T \to \infty$. Let $r = m/k$, $m, k \in \mathbb{N}$, $m < k$ be such a rational number. Consider T large enough so that $T > k^2$, and partition the interval $[0,T]$ into subintervals

$$C_i = \left[(T+\sqrt{T})\frac{i}{k},\, (T+\sqrt{T})\frac{i+1}{k} - \sqrt{T}\right],\ i = 0,1,\dots,k-1\,,$$

$$D_i = \left[(T+\sqrt{T})\frac{i}{k} - \sqrt{T},\, (T+\sqrt{T})\frac{i}{k}\right],\ i = 1,\dots,k-1\,,$$

and observe that by (4.2),

$$P\Big(\tau_{\mathbf{X},T} \in \bigcup_{i=1}^{k-1} D_i\Big) \to 0 \text{ as } T \to \infty.$$

Therefore,

$$P\big(\tau_{\mathbf{X},T} \le rT\big) = P\Big(\max_{0\le i\le m-1} M_{i,T} \ge \max_{m\le i\le k-1} M_{i,T}\Big) + o(1) \tag{4.3}$$

as $T \to \infty$, where $M_{i,T} = \sup_{t\in C_i} X(t)$, $i = 0,1,\dots.k-1$.

Let φ be the function given in Assumption TailSM. Then

$$P\Big(\max_{0\le i\le m-1} M_{i,T} \ge \max_{m\le i\le k-1} M_{i,T}\Big) \tag{4.4}$$

$$= P\Big(\max_{0\le i\le m-1} V_{i,T} \ge \max_{m\le i\le k-1} V_{i,T}\Big) + o(1)\,,$$

where $V_{i,T} = \sup_{t\in C_i} X(t)\mathbf{1}\big(X(t) > \varphi(\sqrt{T})\big)$, $i = 0,1,\dots.k-1$.

Denote by G_T the distribution function of each one of the random variables $V_{i,T}$, and let $W_{i,T} = G_T(V_{i,T})$, $i = 0,1,\dots,k-1$. It is clear that

$$P\Big(\max_{0\le i\le m-1} V_{i,T} \ge \max_{m\le i\le k-1} V_{i,T}\Big) \tag{4.5}$$

$$= P\Big(\max_{0\le i\le m-1} W_{i,T} \ge \max_{m\le i\le k-1} W_{i,T}\Big)\,.$$

Notice, further, that by Assumption TailSM, for every $0 < w_i < 1$, $i = 0, 1, \ldots, k-1$,

$$\lim_{T\to\infty}\Big|P\Big(W_{i,T} \le w_i,\, i = 0, 1, \ldots, k-1\Big) - \prod_{i=0}^{k-1} P\Big(W_{i,T} \le w_i\Big)\Big| = 0\,. \tag{4.6}$$

Let

$$D(T) = \sup_{x\in\mathbb{R}} P\big(\sup_{t\in C_0} X(t) = x\big) + P\big(\sup_{t\in C_0} X(t) \le \varphi(\sqrt{T})\big)\,.$$

By Assumption A, $D(T) \to 0$ as $T \to \infty$. Since for every $0 < w < 1$,

$$w - D(T) \le P\Big(W_{0,T} \le w\Big) \le w\,,$$

we conclude by (4.6) that the law of the random vector $\big(W_{0,T}, \ldots, W_{k-1,T}\big)$ converges weakly, as $T \to \infty$, to the law of a random vector $(U_0, \ldots, U_{k-1})$ with independent standard uniform components. Since this limiting law does not charge the boundary of the set $\{(w_0, w_1, \ldots, w_{k-1}) : \max_{0\le i\le m-1} w_i \le \max_{m\le i\le k-1} w_i\}$, we conclude by (4.3), (4.4) and (4.5) that

$$P\big(\tau_{\mathbf{X},T} \le rT\big) \to P\big(\max_{0\le i\le m-1} U_i \ge \max_{m\le i\le k-1} U_i\big) = m/k = r\,,$$

and so we have established the weak convergence claim of the theorem.

We now prove the uniform convergence of the densities in (4.1). Suppose that the latter fails for some $0 < \varepsilon < 1/2$. There are two possibilities. Suppose first that there is $\theta > 0$, a sequence $T_n \to \infty$ and a sequence $t_n \in [\varepsilon, 1-\varepsilon]$ such that for every n, $T_n f_{\mathbf{X},T_n}(t_n T_n) \ge 1 + \theta$. By compactness we may assume that $t_n \to t_* \in [\varepsilon, 1-\varepsilon]$ as $n \to \infty$. By Lemma 2.1 and the regularity properties of the density, for every n and every $0 < \tau, \delta < 1$ such that

$$\big(1 - (1-\tau)/t_n\big)_+ < \delta < \min\big(\tau/t_n, 1\big) \tag{4.7}$$

we have

$$T_n f_{\mathbf{X},(1-\tau)T_n}(t_n(1-\delta)T_n) \ge T_n f_{\mathbf{X},T_n}(t_n T_n) \ge 1 + \theta\,.$$

Since $t_n \to t_*$, there is a choice of $0 < \tau < 1$ such that

$$1 + \theta > \frac{1}{1-\tau} \tag{4.8}$$

and, moreover, the range in (4.7) is nonempty for all n large enough. Furthermore, we can find $0 < a < b < 1$ such that

$$\big(1 - (1-\tau)/t_n\big)_+ < a < b < \min\big(\tau/t_n, 1\big)$$

for all n large enough. Therefore, for such n

$$(1+\theta)(b-a) \le \int_a^b T_n f_{\mathbf{X},(1-\tau)T_n}(t_n(1-\delta)T_n)\,d\delta$$

$$= \frac{1}{t_n} P\Big(\tau_{\mathbf{X},(1-\tau)T_n} \in \big((1-b)t_n T_n,\, (1-a)t_n T_n\big)\Big) \to \frac{1}{1-\tau}(b-a)$$

as $n \to \infty$ by the already established weak convergence. This contradicts the choice (4.8) of τ.

The second way (4.1) can fail is that there is $0 < \theta < 1$, a sequence $T_n \to \infty$ and a sequence $t_n \in [\varepsilon, 1-\varepsilon]$ such that for every n, $T_n f_{\mathbf{X},T_n}(t_n T_n) \le 1 - \theta$. We can show that this option is impossible as well by appealing, once again, to Lemma 2.1 and using an argument nearly identical to the one described above. Therefore, (4.1) holds, and the proof of the theorem is complete. □

The following corollary is an immediate conclusion of Theorem 4.1. It shows the uniformity of the limiting conditional distribution of the location of the supremum given that it belongs to a suitable subinterval of $[0,T]$.

Corollary 4.2. *Let $\mathbf{X} = (X(t), t \in \mathbb{R})$ be a stationary sample upper semi-continous process, satisfying Assumption TailSM and Assumption A. Let $0 < a_T \le a'_T < b'_T \le b_T < T$ be such that*

$$\liminf_{T\to\infty} \frac{a_T}{T} > 0, \quad \limsup_{T\to\infty} \frac{b_T}{T} < 1, \quad \lim_{T\to\infty} \frac{b'_T - a'_T}{b_T - a_T} = \theta.$$

Then

$$\lim_{T\to\infty} P\Big(\tau_{\mathbf{X},T} \in \big(a'_T, b'_T\big) \Big| \tau_{\mathbf{X},T} \in \big(a_T, b_T\big)\Big) = \theta\,.$$

References

[1] P. Billingsley, *Convergence of probability measures*, second ed., Wiley, New York, 1999. MR-1700749

[2] T. Byczkowski and K. Samotij, *Absolute continuity of stable seminorms*, The Annals of Probability **14** (1986), 289–312. MR-0815972

[3] J. Kim and D. Pollard, *Cube root asymptotics*, The Annals of Statistics **18** (1990), 191–219. MR-1041391

[4] A. Kolmogorov and Y. Rozanov, *On a strong mixing condition for a stationary gaussian process*, Teoria Veroyatnostei i Primeneniya **5** (1960), 222–227. MR-0133175

[5] M.R. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and related properties of random sequences and processes*, Springer Verlag, New York, 1983. MR-0691492

[6] M. Rosenblatt, *Random processes*, Oxford University Press, New York, 1962. MR-0133862

[7] G. Samorodnitsky and Y. Shen, *Is the location of the supremum of a stationary process nearly uniformly distributed?*, Technical report, 2011.

[8] N.D. Ylvisaker, *The expected number of zeros of a stationary gaussian process*, Annals of Mathematical Statistics **36** (1965), 1043–1046. MR-0177458

Acknowledgments. We would like to thank the anonymous referee for the useful comments that helped us to improve the presentation of the material in the paper.