A detailed measure-theoretic proof of Lemma 4.1:

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space in which all the random variables $X_1, X_2, \ldots$ have been defined, and write $\sigma(Y_i)$, $\sigma(Y_i, Y_j)$, etc., for the smallest σ-field generated by Y_i, (Y_i, Y_j), etc.
(i) For fixed $y \in \mathbb{R}$, consider the function $g(Y_j; y) = F(y)I(Y_j \leqslant y)$, which, as a function of Y_j, it is obviously $\sigma(Y_j)$-measurable. If we show that

$$\int_A g(Y_j(\omega); y)\ d\mathbb{P}(\omega) = \int_A I(Y_{j+1}(\omega) \leqslant y)\ d\mathbb{P}(\omega) \tag{1}$$

for arbitrary $A \in \sigma(Y_j)$, it will follows that $g(Y_j; y)$ is indeed a version of the conditional probability $\mathbb{E}[I(Y_{j+1} \leqslant y)|\ \sigma(Y_j)] = \mathbb{P}[Y_{j+1} \leqslant y|\ \sigma(Y_j)]$, which is just an equivalent definition for the conditional probability $\mathbb{P}[Y_{i+1} \leqslant y|\ Y_j]$. In order to show (1), just note that A lies in $\sigma(Y_j)$ if and only if $A = Y_j^{-1}(B)$ for some Borel set B; thus $I(\omega \in A) = I(Y_j \in B)$, so that (1) is equivalent to

$$\mathbb{E}[g(Y_j; y)I(Y_j \in B)] = \mathbb{P}[Y_{j+1} \leqslant y, Y_j \in B] \quad \text{for each Borel set } B. \tag{2}$$

Using the independence of X_{j+1} and $Y_j = \max\{X_1, X_2 \ldots, X_j\}$, the LHS of (2) is

$$\begin{aligned} F(y)\mathbb{E}[I(Y_j \leqslant y)I(Y_j \in B)] &= \mathbb{P}[X_{j+1} \leqslant y]\mathbb{P}[Y_j \leqslant y, Y_j \in B] \\ &= \mathbb{P}[X_{j+1} \leqslant y, Y_j \leqslant y, Y_j \in B], \end{aligned}$$

which equals to the RHS of (2), because $\{X_{j+1} \leqslant y\} \cap \{Y_j \leqslant y\} = \{Y_{j+1} \leqslant y\}$.

Assume now that $i + 1 < j$. It is easily seen that the same function $g(Y_j; y)$, which is $\sigma(Y_{i+1}, Y_j)$-measurable since $\sigma(Y_{i+1}, Y_j)$ contains $\sigma(Y_j)$, satisfies (1) for all A lying in the larger class $\sigma(Y_{i+1}, Y_j)$. Indeed, this fact is equivalent to

$$\mathbb{E}[g(Y_j; y)I((Y_{i+1}, Y_j) \in B)] = \mathbb{P}[Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B]$$

for each Borel set $B \subseteq \mathbb{R}^2$, and the last identity can be easily verified by using the same arguments. Thus, the desired particular case of the Markovian property holds.
(ii) Let h_{i+1} be the function given by (4.4) of the paper, i.e., $h_{i+1}(x) = \frac{dF^{i+1}(x)}{dF(x)}$. Clearly, h_{i+1} is non-decreasing and if $\alpha(F) = \inf\{x : F(x) > 0\}$ is finite then $h_{i+1}(a(F)) > 0$ (see the proof of Theorem 4.2 in the paper). Since $Y_{i+1} \geqslant \alpha(F)$ w.p. 1, we can define the function

$$h(Y_{i+1}; x) = I(Y_{i+1} \leqslant x) + I(Y_{i+1} > x)\frac{F^i(x)}{h_{i+1}(Y_{i+1})},$$

which, for fixed $x \in \mathbb{R}$, is a function of Y_{i+1}, and thus, $\sigma(Y_{i+1})$-measurable. As in part (i), we have to show that

$$\int_A h(Y_{i+1}(\omega); x)\ d\mathbb{P}(\omega) = \int_A I(Y_i(\omega) \leqslant x)\ d\mathbb{P}(\omega)$$

for arbitrary $A \in \sigma(Y_{i+1})$, which is equivalent to

$$\mathbb{E}[h(Y_{i+1};x)I(Y_{i+1} \in B)] = \mathbb{P}[Y_i \leqslant x, Y_{i+1} \in B] \quad \text{for each Borel set } B. \tag{3}$$

Since the d.f. of Y_{i+1} is F^{i+1}, the LHS of (3) equals to

$$\begin{aligned}
&\mathbb{P}[Y_{i+1} \leqslant x, Y_{i+1} \in B] + F^i(x) \int_{(x,\infty)} \frac{I(t \in B)}{h_{i+1}(t)}\, dF^{i+1}(t) \\
&\quad = \mathbb{P}[Y_{i+1} \leqslant x, Y_{i+1} \in B] + F^i(x) \int_{(x,\infty)} I(t \in B)\, dF(t) \\
&\quad = \mathbb{P}[Y_{i+1} \leqslant x, Y_{i+1} \in B] + \mathbb{P}[Y_i \leqslant x]\mathbb{P}[X_{i+1} > x, X_{i+1} \in B] \\
&\quad = \mathbb{P}[Y_{i+1} \leqslant x, Y_{i+1} \in B] + \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, X_{i+1} \in B] \\
&\quad = \mathbb{P}[Y_{i+1} \leqslant x, Y_{i+1} \in B] + \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, Y_{i+1} \in B] \\
&\quad = \mathbb{P}[(\{Y_{i+1} \leqslant x\} \cup \{Y_i \leqslant x, X_{i+1} > x\}) \cap \{Y_{i+1} \in B\}],
\end{aligned}$$

which is the RHS of (3), since $\{Y_{i+1} \leqslant x\} \cup \{Y_i \leqslant x, X_{i+1} > x\} = \{Y_i \leqslant x\}$.

Assume now that $i+1 < j$. As before, for the desired particular case of the reverse Markovian property it suffices to verify the identity

$$\mathbb{E}[h(Y_{i+1};x)I((Y_{i+1}, Y_j) \in B)] = \mathbb{P}[Y_i \leqslant x, (Y_{i+1}, Y_j) \in B]$$

for each Borel set $B \subseteq \mathbb{R}^2$. By the definition of h,

$$\mathbb{E}[h(Y_{i+1};x)I((Y_{i+1}, Y_j) \in B)] = C_1 + C_2,$$

where $C_1 = \mathbb{P}[Y_{i+1} \leqslant x, (Y_{i+1}, Y_j) \in B]$ and

$$\begin{aligned}
C_2 &= F^i(x)\mathbb{E}\left[I(Y_{i+1} > x)\frac{I((Y_{i+1}, Y_j) \in B)}{h_{i+1}(Y_{i+1})}\right] \\
&= F^i(x)\mathbb{E}\left[I(Y_{i+1} > x)\frac{I((Y_{i+1}, \max\{Y_{i+1}, X_{i+2}, \ldots, X_j\}) \in B)}{h_{i+1}(Y_{i+1})}\right] \\
&= F^i(x) \int_{(x,\infty)} \frac{1}{h_{i+1}(t)} \int_{-\infty}^{\infty} I((t, \max\{t,u\}) \in B)\, dF^{j-i-1}(u)\, dF^{i+1}(t) \\
&= F^i(x) \int_{(x,\infty)} \int_{-\infty}^{\infty} I((t, \max\{t,u\}) \in B)\, dF^{j-i-1}(u)\, dF(t) \\
&= F^i(x)\mathbb{P}[X_{i+1} > x, (X_{i+1}, \max\{X_{i+1}, X_{i+2}, \ldots, X_j\}) \in B] \\
&= \mathbb{P}[Y_i \leqslant x]\mathbb{P}[X_{i+1} > x, (X_{i+1}, \max\{X_{i+1}, X_{i+2}, \ldots, X_j\}) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, (X_{i+1}, \max\{X_{i+1}, X_{i+2}, \ldots, X_j\}) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, (Y_{i+1}, Y_j) \in B].
\end{aligned}$$

The above derivation shows that

$$\begin{aligned}
C_1 + C_2 &= \mathbb{P}[Y_{i+1} \leqslant x, (Y_{i+1}, Y_j) \in B] + \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, (Y_{i+1}, Y_j) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, (Y_{i+1}, Y_j) \in B],
\end{aligned}$$

and the desired particular case of the reverse Markovian property follows.

(iii) Using the results in part (i) and (ii), it suffices to verify that for every $x, y \in \mathbb{R}$, the $\sigma(Y_{i+1}, Y_j)$ measurable function $h(Y_{i+1}; x)g(Y_j; y)$ (with g and h as defined above), is a version of the conditional probability $\mathbb{P}[Y_i \leqslant x, Y_{j+1} \leqslant y | \, \sigma(Y_{i+1}, Y_j)]$. Thus, it suffices to verify the identity

$$\mathbb{E}[h(Y_{i+1}; x)g(Y_j; y)I((Y_{i+1}, Y_j) \in B)] = \mathbb{P}[Y_i \leqslant x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B] \quad (4)$$

for all Borel subsets $B \subseteq \mathbb{R}^2$. Observe that if $x \geqslant y$, the relation $Y_{i+1} \leqslant Y_j$ implies that

$$\{g(Y_j; y) \neq 0\} = \{g(Y_j; y) > 0\} \subseteq \{Y_j \leqslant y\} \subseteq \{Y_{i+1} \leqslant x\} \subseteq \{h(Y_{i+1}; x) = 1\};$$

this shows that $h(Y_{i+1}; x)g(Y_j; y) = g(Y_j; y)$, and the LHS of (4), due to (i), is reduced to $\mathbb{E}[g(Y_j; y)I((Y_{i+1}, Y_j) \in B)] = \mathbb{P}[Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B]$, which verifies (4), because $\{Y_{j+1} \leqslant y\} \subseteq \{Y_i \leqslant x\}$. It remains to check (4) for $x < y$. Write

$$\mathbb{E}[h(Y_{i+1}; x)g(Y_j; y)I((Y_{i+1}, Y_j) \in B)] = C_1 + C_2,$$

where

$$\begin{aligned} C_1 &= F(y)\mathbb{P}[Y_{i+1} \leqslant x, Y_j \leqslant y, (Y_{i+1}, Y_j) \in B] \\ &= \mathbb{P}[X_{j+1} \leqslant y]\mathbb{P}[Y_{i+1} \leqslant x, Y_j \leqslant y, (Y_{i+1}, Y_j) \in B] \\ &= \mathbb{P}[X_{j+1} \leqslant y, Y_{i+1} \leqslant x, Y_j \leqslant y, (Y_{i+1}, Y_j) \in B] \\ &= \mathbb{P}[Y_{i+1} \leqslant x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B], \end{aligned}$$

and

$$C_2 = F^i(x)F(y)\mathbb{E}\left[I(Y_{i+1} > x)I(Y_j \leqslant y)\frac{I((Y_{i+1}, Y_j) \in B)}{h_{i+1}(Y_{i+1})}\right].$$

Observe that $Y_j = \max\{Y_{i+1}, W\}$, with $W = \max\{X_{i+2}, \ldots, X_j\}$ if $i + 1 < j$, and $W = -\infty$ if $i + 1 = j$, so that Y_{i+1} and W are independent, and also X_{i+1} and W are independent. Moreover, $Y_j = \max\{X_{i+1}, W\}$ and $X_{i+1} = Y_{i+1}$ when the event $\{Y_i \leqslant x, X_{i+1} > x\}$ occurs. Therefore, we have

$$\begin{aligned} C_2 &= F^i(x)F(y) \\ &\quad \times \mathbb{E}\left[I(Y_{i+1} > x)I(\max\{Y_{i+1}, W\} \leqslant y)\frac{I((Y_{i+1}, \max\{Y_{i+1}, W\}) \in B)}{h_{i+1}(Y_{i+1})}\right] \\ &= F^i(x)F(y) \\ &\quad \times \int_{(x,\infty)} \frac{1}{h_{i+1}(t)} \mathbb{E}[I(\max\{t, W\} \leqslant y)I((t, \max\{t, W\}) \in B)] \, dF^{i+1}(t) \\ &= F^i(x)F(y) \int_{(x,\infty)} \mathbb{E}[I(\max\{t, W\} \leqslant y)I((t, \max\{t, W\}) \in B)] \, dF(t) \\ &= F^i(x)F(y)\mathbb{P}[X_{i+1} > x, \max\{X_{i+1}, W\} \leqslant y, (X_{i+1}, \max\{X_{i+1}, W\}) \in B] \\ &= \mathbb{P}[Y_i \leqslant x]\mathbb{P}[X_{j+1} \leqslant y] \\ &\quad \times \mathbb{P}[X_{i+1} > x, \max\{X_{i+1}, W\} \leqslant y, (X_{i+1}, \max\{X_{i+1}, W\}) \in B] \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}[Y_i \leqslant x, X_{j+1} \leqslant y, X_{i+1} > x, \max\{X_{i+1}, W\} \leqslant y, \\
&\qquad\qquad\qquad\qquad\qquad\qquad (X_{i+1}, \max\{X_{i+1}, W\}) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, X_{j+1} \leqslant y, X_{i+1} > x, Y_j \leqslant y, (Y_{i+1}, Y_j) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B].
\end{aligned}
$$

The above derivation shows that

$$
\begin{aligned}
C_1 + C_2 &= \mathbb{P}[Y_{i+1} \leqslant x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B] \\
&\qquad + \mathbb{P}[Y_i \leqslant x, X_{i+1} > x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B] \\
&= \mathbb{P}[Y_i \leqslant x, Y_{j+1} \leqslant y, (Y_{i+1}, Y_j) \in B],
\end{aligned}
$$

and (4) follows.

Some Comments: (i) It is known that, due to the Markovian character of the Extremal Process, the assertions of Lemma 4.1 hold in a more general manner; for instance, it is clear that for any integers $i_1 < \cdots < i_k < s_1 < \cdots < s_m < j_1 < \cdots < j_r$, the random vectors $\boldsymbol{Y}_1 = (Y_{i_1}, \ldots, Y_{i_k})'$ and $\boldsymbol{Y}_3 = (Y_{j_1}, \ldots, Y_{j_r})'$ are conditionally independent, given $\boldsymbol{Y}_2 = (Y_{s_1}, \ldots, Y_{s_m})'$, and also that the d.f. of $\boldsymbol{Y}_3$ given $\boldsymbol{Y}_2$ depends only on Y_{s_m}, while the d.f. of $\boldsymbol{Y}_1$ given $\boldsymbol{Y}_2$ depends only on Y_{s_1}. Such interesting properties are beyond the scope of the present article, and are not even stated there.
(ii) The assertions of Lemma 4.1 can be rewritten in an informal, less rigorous–more illustrative, way. For example, part (ii) says that for every t_1 in the support S of F (where $S = \{x \in \mathbb{R} : F(x+\epsilon) - F(x-\epsilon) > 0 \text{ for all } \epsilon > 0\}$),

$$
\mathbb{P}(Y_i \leqslant x |\ Y_{i+1} = t_1) = \begin{cases} \dfrac{F^i(x)}{h_{i+1}(t_1)}, & \text{if} \quad x < t_1, \\ 1, & \text{if} \quad x \geqslant t_1, \end{cases}
$$

and if $i + 1 < j$, then

$$
\mathbb{P}(Y_i \leqslant x |\ Y_{i+1} = t_1, Y_j = t_2) = \mathbb{P}(Y_i \leqslant x |\ Y_{i+1} = t_1), \quad x \in \mathbb{R},
$$

holds for all $t_1 \in S$, $t_2 \in S$, with $t_1 \leqslant t_2$. Similarly, part (iii) verifies that for arbitrary $t_1 \leqslant t_2$, $t_1 \in S$, $t_2 \in S$ (where, of course, $t_1 = t_2$ if $i + 1 = j$), the random variables Y_i and Y_{j+1} are conditionally independent, given the event $\{Y_{i+1} = t_1, Y_j = t_2\}$; part (i) can also be written in a similar manner.