Supplementary material for ’Approximate Bayesian Computation of reduced-bias extreme risk measures from heavy-tailed distributions’

Section A and Section B provide additional illustrations and numerical results associated with the experiments on simulated data (Section 4 of the main paper). The proofs of the theoretical results are given in Section C.

A Additional figures

The behavior of $k \mapsto \text{Bias}_k(\hat{U}(1000))$ and $k \mapsto \text{RMSE}_k(\hat{U}(1000))$ associated with the five estimators described in Section 4.2 is depicted on the first two rows of Figure A1 (RPD and Burr distributions with $\gamma = -\rho = 1/2$) and Figure A2 (Fisher and GPD distributions with $\gamma = -\rho = 1/2$). The third row displays the ABC estimator as a function of $k \in \{2, \dots, n-1 = 499\}$ and the associated 90% credible interval computed on a single replication.

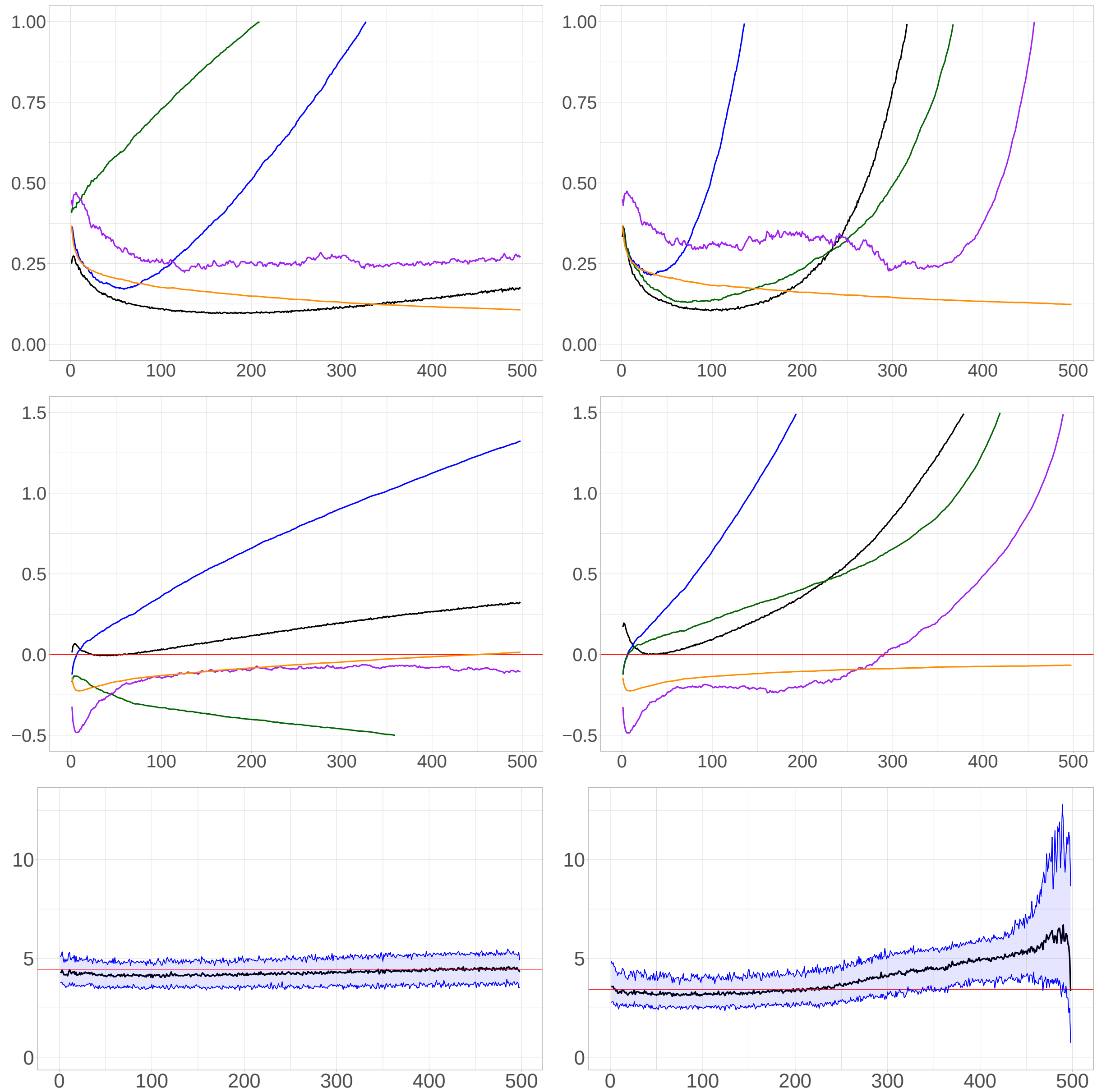

Fig. A1: Illustration on simulated data sets of size $n = 500$ from a RPD (left panel) and a Burr distribution (right panel) with $\gamma = -\rho = 1/2$ in both cases. Top: $k \in \{2, \ldots, n-1\} \mapsto \text{RMSE}_k(\hat{U}(1/p_n = 1000))$ and center: $k \in \{2, \ldots, n-1\} \mapsto \text{Bias}_k(\hat{U}(1/p_n = 1000))$ computed on $N = 500$ replications associated with Weissman (blue), ABC (black), CW (green), PWM (purple) and GPD (orange) estimators. Bottom (in log scale): 90% credible intervals (blue) associated with the ABC estimator (black) computed on one replication. The theoretical extreme quantile $U(1/p_n = 1000)$ is depicted by a red horizontal line.

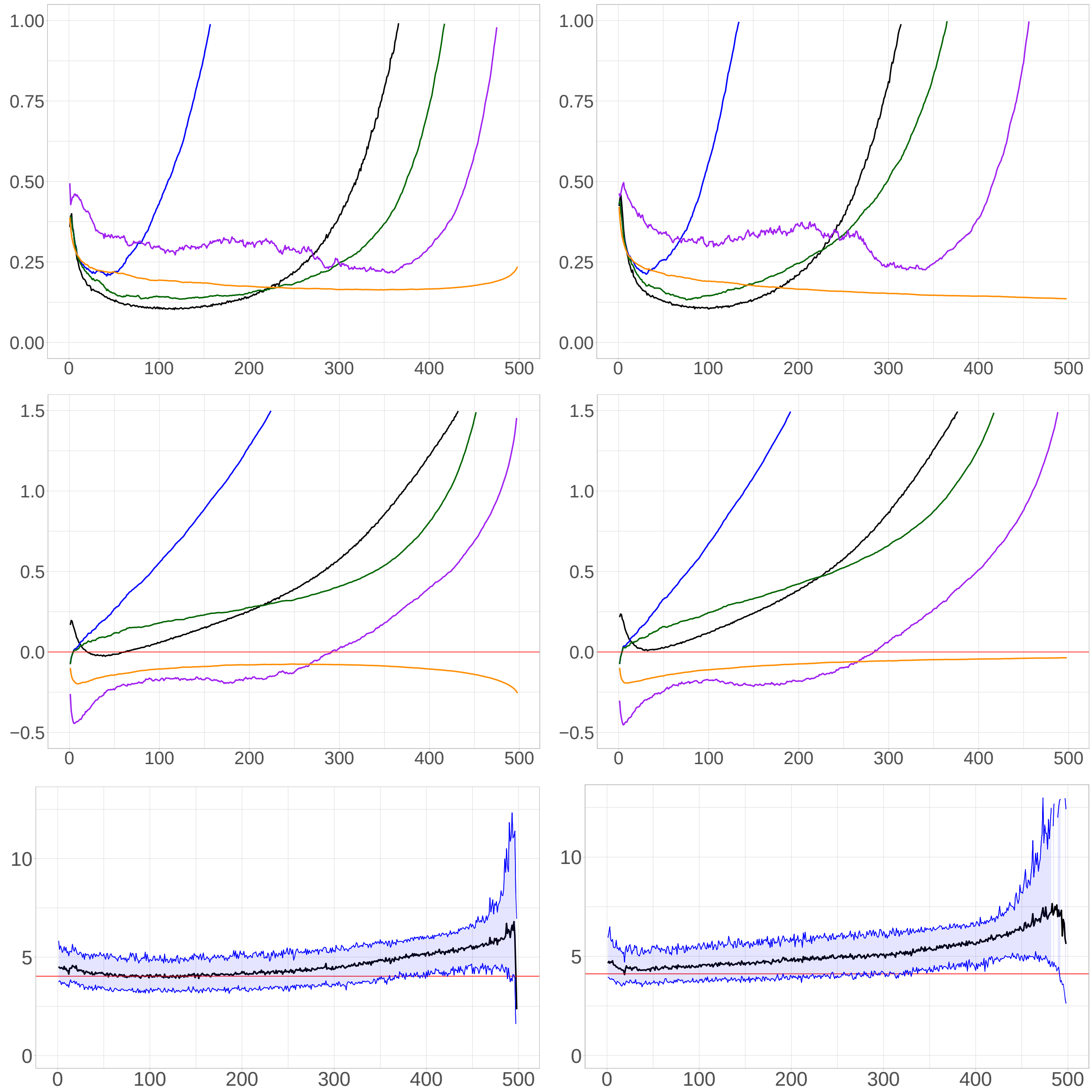

Fig. A2: Illustration on simulated data sets of size $n = 500$ from a Fisher distribution (left panel) and a GPD (right panel) with $\gamma = -\rho = 1/2$ in both cases. Top: $k \in \{2, \ldots, n-1\} \mapsto \text{RMSE}_k(\hat{U}(1/p_n = 1000))$ and center: $k \in \{2, \ldots, n-1\} \mapsto \text{Bias}_k(\hat{U}(1/p_n = 1000))$ computed on $N = 500$ replications associated with Weissman (blue), ABC (black), CW (green), PWM (purple) and GPD (orange) estimators. Bottom (in log scale): 90% credible intervals (blue) associated with the ABC estimator (black) computed on one replication. The theoretical extreme quantile $U(1/p_n = 1000)$ is depicted by a red horizontal line.

B Additional tables

The M_1RMSEs are provided in Table B1 for the RPD, in Table B2 for the Burr distribution and in Table B3 for Fréchet, Fisher, GPD, Inverse Gamma, and Student t distributions.

Table B1: M_1RMSE associated with five estimators of $\log U(1/p_n = 1000)$ computed on $N = 500$ replications of a data set of size $n = 500$ from a RPD. The best result is emphasized in bold. M_1RMSEs larger than 1 are not reported.

RPD	Weissman	ABC	CW	PWM	GPD
$\gamma = 1/8$					
$\rho = -1/8$	0.0175	**0.0106**	0.0475	0.0202	0.0176
$\rho = -1/4$	0.0165	**0.0092**	0.0385	0.0198	0.0132
$\rho = -1/2$	0.0107	**0.0059**	0.0254	0.0140	0.0113
$\rho = -1$	0.0060	**0.0037**	0.0111	0.0115	0.0138
$\rho = -2$	**0.0035**	0.0042	0.0065	0.0073	0.0129
$\gamma = 1/4$					
$\rho = -1/8$	0.0700	**0.0423**	0.1901	0.0810	0.0572
$\rho = -1/4$	0.0638	**0.0349**	0.1359	0.0740	0.0425
$\rho = -1/2$	0.0427	**0.0237**	0.1016	0.0558	0.0303
$\rho = -1$	0.0240	**0.0150**	0.0445	0.0459	0.0454
$\rho = -2$	**0.0139**	0.0168	0.0261	0.0293	0.0443
$\gamma = 1/2$					
$\rho = -1/8$	0.2801	**0.1691**	0.7603	0.3239	0.2076
$\rho = -1/4$	0.2551	0.1397	0.5435	0.2961	**0.1203**
$\rho = -1/2$	0.1707	**0.0950**	0.4065	0.2234	0.1071
$\rho = -1$	0.0960	**0.0598**	0.1779	0.1835	0.1301
$\rho = -2$	**0.0555**	0.0671	0.1045	0.1171	0.1599
$\gamma = 1$					
$\rho = -1/8$	-	**0.6764**	-	-	-
$\rho = -1/4$	-	**0.5588**	-	-	-
$\rho = -1/2$	0.6830	**0.3799**	-	0.8935	-
$\rho = -1$	0.3840	**0.2393**	0.7117	0.7340	0.8531
$\rho = -2$	**0.2219**	0.2685	0.4179	0.4683	0.8102

Table B2: M_1RMSE associated with five estimators of $\log U(1/p_n = 1000)$ computed on $N = 500$ replications of a data set of size $n = 500$ from a Burr distribution. The best result is emphasized in bold. M_1RMSEs larger than 1 are not reported.

Burr	Weissman	ABC	CW	PWM	GPD
$\gamma = 1/8$					
$\rho = -1/8$	0.0593	0.0563	0.0585	0.0775	**0.0264**
$\rho = -1/4$	0.0268	**0.0182**	0.0249	0.0321	0.0246
$\rho = -1/2$	0.0134	**0.0065**	0.0081	0.0142	0.0173
$\rho = -1$	0.0062	**0.0029**	0.0049	0.0056	0.0152
$\rho = -2$	0.0035	**0.0027**	0.0055	0.0032	0.0132
$\gamma = 1/4$					
$\rho = -1/8$	0.2373	0.2251	0.2339	0.3101	**0.0961**
$\rho = -1/4$	0.1071	0.0727	0.0995	0.1284	**0.0427**
$\rho = -1/2$	0.0537	**0.0259**	0.0324	0.0568	0.0566
$\rho = -1$	0.0249	**0.0115**	0.0194	0.0225	0.0520
$\rho = -2$	0.0146	**0.0103**	0.0198	0.0128	0.0456
$\gamma = 1/2$					
$\rho = -1/8$	0.9493	0.9005	0.9355	-	**0.8022**
$\rho = -1/4$	0.4286	0.2908	0.3978	0.5136	**0.0626**
$\rho = -1/2$	0.2147	**0.1034**	0.1297	0.2272	0.1237
$\rho = -1$	0.0994	**0.0460**	0.0778	0.0900	0.1756
$\rho = -2$	0.0556	**0.0431**	0.0883	0.0513	0.1632
$\gamma = 1$					
$\rho = -1/8$	-	-	-	-	-
$\rho = -1/4$	-	-	-	-	-
$\rho = -1/2$	0.8559	**0.4137**	0.5190	0.9090	-
$\rho = -1$	0.3978	**0.1841**	0.3112	0.3602	0.8302
$\rho = -2$	0.2263	**0.1723**	0.3522	0.2052	0.8257

Table B3: M_1RMSE associated with five estimators of $\log U(1/p_n = 1000)$ computed on $N = 500$ replications of data sets of size $n = 500$ from five heavy-tailed distributions. The best result is emphasized in bold.

	Weissman	ABC	CW	PWM	GPD
Fréchet ($\rho = -1$)					
$\gamma = 1/8$	0.0047	**0.0031**	0.0047	0.0042	0.0135
$\gamma = 1/4$	0.0189	**0.0122**	0.0186	0.0169	0.0460
$\gamma = 1/2$	0.0759	**0.0488**	0.0745	0.0679	0.1624
$\gamma = 1$	0.3035	**0.1953**	0.2981	0.2716	0.8276
Fisher ($\rho = -\gamma$)					
$\gamma = 1/8$	0.0495	0.0412	0.0490	0.0683	**0.0329**
$\gamma = 1/4$	0.0860	0.0562	0.0766	0.1236	**0.0555**
$\gamma = 1/2$	0.2062	**0.1022**	0.1345	0.2158	0.1632
$\gamma = 1$	0.4578	**0.2051**	0.3587	0.3849	0.8800
GPD ($\rho = -\gamma$)					
$\gamma = 1/8$	0.0651	0.0551	0.0612	0.0785	**0.0253**
$\gamma = 1/4$	0.1088	0.0729	0.0977	0.1362	**0.0441**
$\gamma = 1/2$	0.2118	**0.1042**	0.1322	0.2253	0.1357
$\gamma = 1$	0.4030	**0.1823**	0.3030	0.3492	0.7971
Inverse Gamma ($\rho = -\gamma$)					
$\gamma = 1/8$	0.0256	**0.0146**	0.0234	0.0305	0.0260
$\gamma = 1/4$	0.0552	**0.0272**	0.0444	0.0579	0.0529
$\gamma = 1/2$	0.1268	**0.0605**	0.0843	0.1217	0.1755
$\gamma = 1$	0.3082	**0.1789**	0.2739	0.2529	0.81.98
Student t ($\rho = -2\gamma$)					
$\gamma = 1/8$	0.0303	**0.0263**	0.0290	0.0427	0.0272
$\gamma = 1/4$	0.0572	**0.0343**	0.0477	0.0819	0.0541
$\gamma = 1/2$	0.1406	0.0619	**0.0501**	0.1170	0.1601
$\gamma = 1$	0.2721	**0.1579**	0.4192	0.2192	0.7659

C Proofs

Proof of Proposition 1.

Let us introduce $\varphi(x) = \log U(\exp x)$ for all $x \geq 0$. Replacing in (8) yields the equation

$$\frac{1}{A(t)}(\varphi(\log t + \log y) - \varphi(\log t) - \gamma \log y) = K_\rho(y),$$

for all $t \geq 1$ and $y > 0$, or equivalently, letting $s = \log t$ and $x = \log y$,

$$\frac{1}{A(\exp s)}\left(\frac{\varphi(s+x) - \varphi(s)}{x} - \gamma\right) = \frac{K_\rho(\exp x)}{x},$$

for all $x \neq 0$ and $s \geq 0$. Letting $x \to 0$ and remarking that $K_\rho(\exp x)/x \to \rho$ yield

$$\varphi'(s) = \gamma + \rho A(\exp s) = \gamma(1 + \beta \exp(\rho s)),$$

in view of (6). Integrating, it follows that, for all $s \geq 0$,

$$\varphi(s) = \varphi(0) + \gamma(s + \beta K_\rho(\exp s)),$$

leading to

$$U(x) = U(1)x^\gamma \exp(\beta\gamma K_\rho(x)),$$

for all $x \geq 1$. Conversely, it is easily checked that the above U function is a solution of (8). Finally, the condition $\beta \geq -1$ is required to ensure that U is increasing. ■

Proof of Proposition 2.

(i) The tail quantile function of Y is given for all $y \geq 1$ by

$$U_\alpha(y) = U(y/\alpha)/U(1/\alpha) \tag{23}$$

so that (5) can be rewritten with $t = 1/\alpha$ as

$$\lim_{\alpha\to 0} \frac{1}{A(1/\alpha)} (\log U_\alpha(y) - \gamma \log(y)) = K_\rho(y),$$

or equivalently,

$$\lim_{\alpha\to 0} \frac{1}{A(1/\alpha)} (\log U_\alpha(y) - \gamma \log(y) - A(1/\alpha)K_\rho(y)) = 0.$$

Taking account of (6), it follows

$$\lim_{\alpha\to 0} \alpha^\rho \left(\log U_\alpha(y) - \gamma \log(y) - \beta\gamma\alpha^{-\rho}K_\rho(y)\right) = 0,$$

leading to

$$\lim_{\alpha\to 0} \alpha^\rho \left(\log U_\alpha(y) - \log U_{\mathrm{RPD}}(y \mid \gamma, \rho, \beta\alpha^{-\rho})\right) = 0,$$

if $U_{\mathrm{RPD}}(1) = 1$. Taking the exponential concludes the proof.

(ii) The result is a consequence of (23) and of the identity

$$\frac{U_{\mathrm{RPD}}(y/\alpha \mid \gamma, \rho, \beta)}{U_{\mathrm{RPD}}(1/\alpha \mid \gamma, \rho, \beta)} = (1/\alpha)^\gamma \exp\left(\frac{\beta\gamma}{\rho}\alpha^{-\rho}(y^\rho - 1)\right) = U_{\mathrm{RPD}}(y \mid \gamma, \rho, \beta\alpha^{-\rho})$$

that holds for all $y \geq 1$. ■

Proof of Proposition 3.
The probability weighted moment of order $a > -1$ is given by

$$m_Y(a) = (a+1)^2\,\mathbb{E}((\log Y)(\bar{F}_{\mathrm{RPD}}(Y))^a) = (a+1)^2 \int_{U_{\mathrm{RPD}}(1)}^{+\infty} \log(x)(\bar{F}_{\mathrm{RPD}}(x))^a f_{\mathrm{RPD}}(x)\,\mathrm{d}x,$$

where $\bar{F}_{\mathrm{RPD}}$ and f_{RPD} denote respectively the survival function and the density function associated with the tail quantile function U_{RPD} given in Definition 1. The change of variable $y \mapsto x = U_{\mathrm{RPD}}(y)$ yields

$$\begin{aligned} m_Y(a) &= (a+1)^2 \int_1^{+\infty} \log(U_{\mathrm{RPD}}(y)) y^{-a-2}\,\mathrm{d}y, \\ &= (a+1)^2 \int_1^{+\infty} \log\left(U_{\mathrm{RPD}}(1) y^{\gamma} \exp(\beta\gamma K_\rho(y))\right) y^{-a-2}\,\mathrm{d}y, \\ &= (a+1)^2 \log\left(U_{\mathrm{RPD}}(1)\right) \int_1^{+\infty} y^{-a-2}\,\mathrm{d}y + (a+1)^2\gamma \int_1^{+\infty} \log(y)\, y^{-a-2}\,\mathrm{d}y \\ &+ (a+1)^2\beta\gamma \int_1^{+\infty} K_\rho(y) y^{-a-2}\,\mathrm{d}y. \end{aligned}$$

The first term vanishes given that $U_{\mathrm{RPD}}(1) = 1$. Integrating by parts the second term concludes the proof. ■

Proof of Proposition 4.
Suppose $U_{\mathrm{RPD}}(1) = 1$, $a\gamma < 1$ and introduce $c = -a\beta\gamma/\rho$. Then, a direct calculation yields

$$\begin{aligned} \mathrm{CTM}_Y(a,\alpha) &= \frac{1}{\alpha}\int_0^{\alpha} U_{\mathrm{RPD}}^a(1/v \mid \gamma,\rho,\beta)\mathrm{d}v \\ &= \frac{1}{\alpha}\int_0^{\alpha} v^{-a\gamma}\exp(a\beta\gamma K_\rho(1/v))\mathrm{d}v \\ &= \frac{\exp(c)}{\alpha}\int_0^{\alpha} v^{-a\gamma}\exp(-cv^{-\rho})\mathrm{d}v \\ &= \exp(c)\alpha^{-a\gamma}\int_0^1 u^{-a\gamma}\exp(-c\alpha^{-\rho}u^{-\rho})\mathrm{d}u, \end{aligned}$$

which is the desired result. Finally, note that, if $\beta > 0$, then $c > 0$ too and the CTM can be computed as

$$\mathrm{CTM}_Y(a,\alpha) = -\frac{\exp(c)}{\alpha\rho} c^{\frac{1-a\gamma}{\rho}} \Gamma_\ell\left(\frac{a\gamma - 1}{\rho}, c\alpha^{-\rho}\right),$$

where $\Gamma_\ell(\cdot,\cdot)$ is the incomplete lower gamma function. ■

Proof of Lemma 1.
The result is a straightforward consequence of (23) in the proof of Proposition 2: Letting $\alpha = k/n$ and $y = d_n = k/(np_n)$ yields $U_{k/n}(d_n) = U(1/p_n)/U(n/k)$ and the result is proved. ■

Proof of Theorem 5.
Let $d_n = k/(np_n)$ be the extrapolation factor. The following expansion holds:

$$
\begin{aligned}
\log \hat{U}(1/p_n) - \log U(1/p_n) &= (\log \hat{U}(1/p_n) - \overline{\log U}(1/p_n)) + (\overline{\log U}(1/p_n) - \log U(1/p_n)) \\
&= A_{1,n} + A_{2,n} + A_{3,n} + B_n, \\
\text{with } A_{1,n} &= \log X_{n-k,n} - \log U(n/k), \\
A_{2,n} &= (\hat{\gamma} - \gamma) \log(d_n), \\
A_{3,n} &= \hat{\gamma} \hat{\beta}_n K_{\hat{\rho}}(d_n) - \gamma \beta_n K_\rho(d_n), \\
B_n &= \log U_{\text{RPD}}(d_n \mid \gamma, \rho, \beta(n/k)^\rho) - \log U_{k/n}(d_n).
\end{aligned}
$$

Each term is considered separately. First, (de Haan and Ferreira, 2006, Theorem 2.4.1) yields

$$
\frac{\sqrt{k_n}}{\log(d_n)} A_{1,n} = O_p(1/\log(d_n)), \tag{24}
$$

since $\sqrt{k} A(n/k) \to \lambda$ as $n \to \infty$. Second, letting $\xi_n := \sqrt{k}(\hat{\gamma} - \gamma)$, one has

$$
\frac{\sqrt{k_n}}{\log(d_n)} A_{2,n} = \xi_n. \tag{25}
$$

Third, let us rewrite $A_{3,n}$ as

$$
A_{3,n} = \beta_n \left(\hat{\gamma} \frac{\hat{\beta}_n}{\beta_n} K_{\hat{\rho}}(d_n) - \gamma K_\rho(d_n) \right) = \gamma \beta (n/k)^\rho \left(K_{\hat{\rho}}(d_n) O_P(1) - K_\rho(d_n) \right),
$$

since $\hat{\gamma} \xrightarrow{\mathbb{P}} \gamma$ and $\hat{\beta}_n/\beta_n = O_P(1)$ by assumption. Remark that $K_\rho(d_n) \to -1/\rho$ as $n \to \infty$ while $|K_{\hat{\rho}}(d_n)| \leq -2/\hat{\rho} \leq -2/\rho^\dagger$ almost surely. As a consequence,

$$
\frac{\sqrt{k_n}}{\log(d_n)} A_{3,n} = O_P(1/\log(d_n)), \tag{26}
$$

under the assumption $\sqrt{k} A(n/k) \to \lambda$ as $n \to \infty$. Finally, B_n is a non-random term controlled with (de Haan and Ferreira, 2006, Eq. (3.2.7)): For all $\varepsilon > 0$, there exists t_0 such that for $t \geq t_0$ and $y \geq 1$,

$$
\left| \frac{1}{A(t)} (\log U(ty) - \log U(t) - \gamma \log(y)) - K_\rho(y) \right| \leq \varepsilon y^{\rho+\varepsilon}.
$$

Taking account of (6) and considering $y = d_n \to \infty$ and $t = n/k \to \infty$, it follows

$$
|\log U(1/p_n) - \log U(n/k) - \gamma \log(d_n) - \gamma \beta (n/k)^\rho K_\rho(d_n)| \leq \varepsilon d_n^{\rho+\varepsilon} |A(n/k)|,
$$

or equivalently, in view of Definition 1 and (23) in the proof of Proposition 2, $|B_n| \leq \varepsilon d_n^{\rho+\varepsilon} |A(n/k)|$, so that

$$
\frac{\sqrt{k_n}}{\log(d_n)} |B_n| \leq \varepsilon |\lambda| \frac{d_n^{\rho+\varepsilon}}{\log(d_n)} (1 + o(1)), \tag{27}
$$

under the assumption $\sqrt{k} A(n/k) \to \lambda$ as $n \to \infty$. Combining (24)–(27), it follows that

$$
\frac{\sqrt{k}}{\log(d_n)} \left(\log \hat{U}(1/p_n) - \log U(1/p_n) \right) = \xi_n + O_P \left(\frac{1}{\log(d_n)} \right) + O \left(\frac{d_n^{\rho+\varepsilon}}{\log(d_n)} \right),
$$

Choosing $\varepsilon < -\rho$ concludes the proof. ■

Proof of Lemma 2.
Recall that

$$\mathrm{CTM}_X(a, p_n) = \frac{1}{p_n} \int_0^{p_n} U^a(1/v)\mathrm{d}v,$$

with $U(1/v) = U(n/k)U_{k/n}(k/(nv))$ from (23) in the proof of Proposition 2. Replacing, one obtains

$$\mathrm{CTM}_X(a, p_n) = \frac{1}{p_n} U^a(n/k) \int_0^{p_n} U^a_{k/n}(k/(nv))\mathrm{d}v = d_n U^a(n/k) \int_0^{1/d_n} U^a_{k/n}(1/u)\mathrm{d}u,$$

with the change of variable $u = nv/k$. Remarking that

$$d_n \int_0^{1/d_n} U^a_{k/n}(1/u)\mathrm{d}u = \mathrm{CTM}_{k/n}(a, 1/d_n)$$

concludes the proof. ■