Extreme Age of Information for Wireless-Powered Communication Systems

Nikolaos I. Miridakis, *Senior Member, IEEE*, Zheng Shi, *Member, IEEE*, Theodoros A. Tsiftsis, *Senior Member, IEEE*, and Guanghua Yang, *Senior Member, IEEE*

arXiv:2201.07437v2 [cs.IT] 20 Jan 2022

Abstract

The extreme or maximum age of information (AoI) is analytically studied for wireless communication systems. In particular, a wireless powered single-antenna source node and a receiver (connected to the power grid) equipped with multiple antennas are considered when operated under independent Rayleigh-faded channels. Via the extreme value theory and its corresponding statistical features, we demonstrate that the extreme AoI converges to the Gumbel distribution whereas its corresponding parameters are obtained in straightforward closed-form expressions. Capitalizing on this result, the risk of the extreme AoI realization is analytically evaluated according to some relevant performance metrics, while some useful engineering insights are manifested.

Index Terms

Age of information (AoI), extreme value theory, low-latency, wireless power transfer.

I. INTRODUCTION

Age of information (AoI) represents an insightful performance metric for wireless communication systems, which indicates the timeliness of data delivery. Doing so, it has numerous applications in a set of various recent and upcoming real-time networking paradigms, e.g., Internet-of-things, vehicular and machine-type communications, healthcare, infrastructure and

Corresponding Author: T. A. Tsiftsis.

The authors are with the School of Intelligent Systems Science & Engineering, Jinan University, Zhuhai Campus, Zhuhai 519070, China. N. I. Miridakis is also with the Dept. of Informatics and Computer Engineering, University of West Attica, Aegaleo 12243, Greece (e-mails: nikozm@uniwa.gr, zhengshi@jnu.edu.cn, theo_tsiftsis@jnu.edu.cn, ghyang@jnu.edu.cn).

environment monitoring, as well as low-latency and delay-critical services [1], [2]. Source nodes in these applications are usually simple sensor devices which are battery-powered and thus become energy-constrained. The energy efficiency can be enhanced, when energy harvesting is implemented at the considered system nodes. In particular, energy harvesting via wireless power transfer (WPT) can be used for enabling perpetual operation of wireless devices. In fact, WPT has attracted increasing interests mainly due to its capability of converting received RF signals into electricity, which in turn is able to provide stable and controllable power to prolong the lifetime of low-power energy-constrained autonomous networks [3].

In WPT systems, AoI plays a pivotal role since the source nodes are in principle power-limited and thus the scenario of erroneous detection (which is accompanied with consecutive retransmissions) dramatically affects the freshness of information at the receiver. To this end, various research studies have analyzed the impact of AoI in such systems; see, e.g., [3]–[5] and relevant references therein. Nonetheless, most studies so far have focused on the analysis and/or optimization of the average AoI (or peak AoI). Although the average AoI defines a key statistical performance metric, it can not capture the effect of information aging for *extremely* rare events. A paradigm of particular interest is the maximum AoI during the communication process of the transceiver, which is located in the right tail of the (instantaneous) AoI distribution. The extreme (or maximum) AoI may influence more drastically the system performance rather than its average counterpart in certain applications; such as autonomous vehicles, tactile Internet as well as other exciting yet delay-critical and low-latency services. To our knowledge, the effect of extreme AoI has only been analytically studied in [6] and [7]. In the latter works, however, the (conditional on some time threshold) extreme AoI was approached by the generalized extreme value distribution. Yet, its corresponding parametric values were numerically computed, while the energy limitation of the supporting source nodes was not considered therein.

Capitalizing on the aforementioned observations, for the first time in this Letter, the (unconditional) extreme AoI is analytically studied for WPT systems. Using the statistical tools of the extreme value theory, we explicitly indicate that the extreme AoI is suitably approached by the Gumbel distribution. The parametric values of the said distribution are derived in simple closed-form expressions. Furthermore, the *risk* arising from an extremely high aging of information is analytically quantified via some relevant performance metrics; namely, the value-at-risk (VaR) and conditional VaR (CVaR). With the latter metrics at hand, useful engineering insights are revealed which may become impactful for the system design of delay-critical and/or low-latency

wireless networking applications.

Notation: $\mathbb{E}[\cdot]$ is the expectation operator; $\Pr[\cdot]$ stands for the probability operator; $|\cdot|$ represents the absolute (scalar) value operator; $y|z$ denotes that y is conditioned on z event; $\hat{x}$ denotes an estimate of x; $\text{li}(\cdot)$ is the logarithmic integral function [8, Eq. (4.211.2)]; $\Gamma(\cdot, \cdot)$ is the upper incomplete Gamma function [8, Eq. (8.350.2)]; ϵ denotes the Euler-Mascheroni constant [8, Eq. (8.367)].

II. SYSTEM MODEL

Consider a single-antenna source node that generates status updates and sends them at a given destination equipped with $N \geq 1$ antennas which is connected to the power grid. The transmitting source has energy harvesting capabilities and fills its battery via WPT from a dedicated energy transmitter (ET); it stores energy in a capacitor of size S and transmits fresh data only when its battery is full. Then, all its harvested energy is used striving for a successful data reception. It is assumed that ET is connected to the power grid, whereas it continuously broadcasts a certain RF signal with a fixed power P_t. Thereupon, assuming a linear WPT model, the stored energy of the source at the lth timeslot is modeled by $E_l \triangleq \min\{E_{l-1} + \eta P_\text{t}|u_l|^2, S\}$ when $E_{l-1} < S$; or $E_l \triangleq \min\{\eta P_\text{t}|u_l|^2, S\}$ when $E_{l-1} \geq S$. Also, η represents the RF-to-DC conversion efficiency, while u_l denotes the channel fading coefficient for the link between the source node and ET at the lth timeslot. Although a sophisticated nonlinear WPT modeling is more appropriate for realistic conditions, the adopted simplified linear model serves as a lower performance bound on the nonlinear harvested energy used in practice [3], [9]. In addition, the energy and communication links are assumed orthogonal (e.g., occupy different frequency bands) so as to avoid mutual interference. Further, the data communication is achieved in consecutive timeslots with a slot size of one time unit.[1] In Fig. 1, the AoI evolution across time is illustrated for the considered WPT system model.

All the included wireless links experience independent quasi-static Rayleigh fading conditions, whereon the channel remains fixed for the duration of one timeslot while it changes afterwards. At the receiver, with a perfect channel-state information (CSI) at hand, maximum ratio combining

---

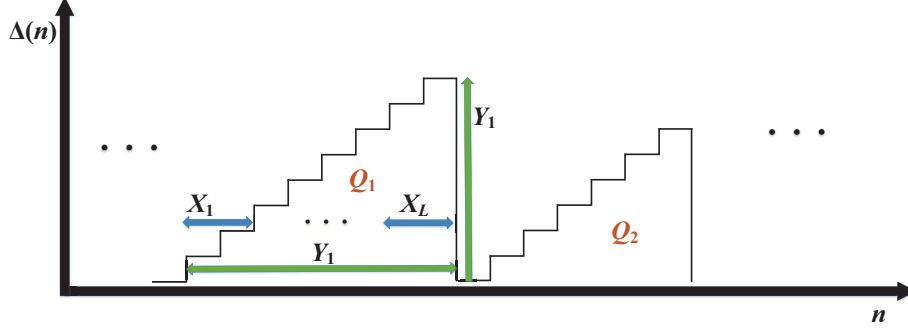[1]Henceforth, the energy and power metrics become equal and will be used interchangeably.

Fig. 1. The AoI progression of the considered system model.

(MRC) is utilized in order to enhance the received channel gains. Consequently, the received signal-to-noise ratio (SNR) at the ith timeslot reads as

$$\gamma_i = \frac{S}{N_0} \sum_{v=1}^{N} |h_{i,v}|^2, \tag{1}$$

where N_0 is the additive white Gaussian noise power and $h_{i,v}$ denotes a zero-mean and unit-variance circularly symmetric complex Gaussian fading coefficient at the ith timeslot and v^{th} received antenna element. In Fig. 1, X_i represents the number of consecutive timeslots required for the capacitor being fully charged. When the energy level reaches the S value, a signal transmission occurs; in the case when a successful reception occurs, $\log_2(1 + \gamma_i) \geq R$ is satisfied with R standing for a predetermined data rate (in bps/Hz) and the AoI is reset to zero. Otherwise, the capacitor starts to recharge in the following timeslots and then it repeats the signal retransmission until a successful transmission occurs. In the latter case, all the timeslots required for a successful reception (i.e., the interarrival time between consecutive decoded messages) are concretely defined as $Y_j \triangleq \sum_{i=1}^{L} X_i$ with $L < \infty$ (as per Fig. 1).

III. PERFORMANCE METRICS

The instantaneous AoI at a given timeslot m is defined as

$$\Delta(m) \triangleq m - \mathcal{G}(m), \tag{2}$$

where $\mathcal{G}(m)$ is the generation time of the last update message.

A. Average AoI

For a *stationary and ergodic* information updating process, the time average AoI tends to its ensemble average counterpart given an asymptotically high number of consecutive samples. Therefore, it holds that

$$\overline{\Delta} \triangleq \frac{1}{M} \sum_{m=1}^{M} \Delta(m) = \frac{K}{M} \frac{1}{K} \sum_{i=1}^{K} Q_i = \frac{\mathbb{E}[Q]}{\mathbb{E}[Y]}, \quad M \to +\infty, \tag{3}$$

where $\mathbb{E}[Q] \triangleq \frac{1}{K} \sum_{i=1}^{K} Q_i$ and $1/\mathbb{E}[Y]$ denotes the steady-state rate of the packet arrival time, such that $\mathbb{E}[Y] \triangleq \lim_{M \to +\infty} \frac{M}{K} = \frac{1}{K} \sum_{i=1}^{K} Y_i$. Also, Q_i denotes the disjoint area in Fig. 1 which is computed as

$$Q_i = \sum_{u=1}^{Y_i} u = \frac{1}{2} \left[Y_i(Y_i + 1) \right]. \tag{4}$$

It turns out that [3, Thm. 1]

$$\overline{\Delta} = \frac{1}{2} \left(\frac{\mathbb{E}[Y^2]}{\mathbb{E}[Y]} + 1 \right) = \frac{1 + 3g + g^2}{2(g+1)} + \frac{(1+g)p}{1-p} + \frac{1}{2}, \tag{5}$$

where

$$p \triangleq \Pr[\log_2(1 + \gamma_i) < R] = 1 - \frac{\Gamma\left(N, \frac{wd_c^{\alpha}(2^R - 1)}{S/N_0}\right)}{(N-1)!} \tag{6}$$

is the outage probability with w, d_c and α denoting the propagation attenuation at a reference distance of 1 m., transceiver distance of the communication link and path-loss exponent, respectively. Also,

$$g \triangleq \frac{wd_e^{\alpha}S}{\eta P_{\text{t}}}, \tag{7}$$

where d_e is the transceiver distance of the WPT link.

For the limiting case when $N \to +\infty$ (i.e., a massive receive antenna array) and with the aid of [10, Eq. (8.11.5)], $p \to 0^+$ and (5) simplifies to $\overline{\Delta}^{(N \to +\infty)} \approx (2 + 4g + g^2)/(2 + 2g)$. Moreover, in the asymptotically high transmission power regime (where $S/P_{\text{t}} \to 0^+$ and thus $g \to 0^+$), the average AoI tends to one timeslot.[2]

---

[2]To this end, the simultaneous energy harvesting and data transmission is required. Doing so, the *energy half-duplex problem* can be avoided by assuming two antennas and two energy storage devices (one used for data transmission and the other for energy harvesting, correspondingly), as in [3] and [9].

B. AoI Variance

Following similar lines of reasoning as for the derivation of (3), we get the AoI variance derived by

$$\mathcal{V}_\Delta = \frac{\mathbb{E}[Q^2]}{\mathbb{E}[Y]} - \overline{\Delta}^2 = \frac{\mathbb{E}\left[[Y(Y+1)]^2\right]}{4\mathbb{E}[Y]} - \frac{\mathbb{E}^2\left[Y(Y+1)\right]}{4\mathbb{E}^2[Y]}$$

$$= \frac{1}{\mathbb{E}[Y]}\left(\frac{\mathbb{E}[Y^4]}{4} + \frac{\mathbb{E}[Y^3]}{2} + \frac{\mathbb{E}[Y^2]}{4}\right) - \frac{(\mathbb{E}[Y^2] + \mathbb{E}[Y])^2}{4\mathbb{E}^2[Y]}. \tag{8}$$

The above formula yields a closed-form expression given by

$$\begin{aligned}
\mathcal{V}_\Delta = & \frac{1}{4(g+1)^2(1-p)^3}\left[g^5\left(p^3 + 11p^2 + 11p + 1\right)\right. \\
& + g^4\left(32p^2 + 78p - 2p^3 + 12\right) \\
& + g^3\left(26p^2 + 174p - 2p^3 + 42\right) \\
& + 2g^2\left(p^3 + 5p^2 + 89p + 25\right) \\
& \left. + 4g\left(3p^2 + 22p + 5\right) + 4p(p+5)\right].
\end{aligned} \tag{9}$$

The proof of (9) is relegated in the Appendix. For the limiting case when $N \to +\infty$, $p \to 0^+$ and (9) simplifies to

$$\mathcal{V}_\Delta^{(N \to +\infty)} \approx \frac{g\left(g^4 + 12g^3 + 42g^2 + 50g + 20\right)}{4(g+1)^2}. \tag{10}$$

Further, in the asymptotically high transmission power regime (where $g \to 0^+$), the AoI variance tends to zero, as expected.

C. Extreme AoI Statistics

Although the AoI variance is a powerful statistical tool and (in contrast to the average AoI) provides useful information on the fluctuation of AoI, it is not sufficient to estimate the behavior of extreme cases (e.g., the maximum attainable AoI). To this end, we resort to the extreme value theory which efficiently describes rare events. Subsequently, we show that the Gumbel distribution suitably models the statistical behavior of extreme (rare) AoI events for the considered system model.

We commence by recognizing that X_i (i.e., the number of consecutive timeslots required for a fully-charged capacitor) follows a Poisson distribution as per (A.1). Then, the jth interarrival time between two successive correctly decoded update messages (say, the instantaneous AoI)

$Y_j = \sum_i^L X_i$ with $L < \infty$ is also a Poisson RV with a corresponding PMF (conditioned on L attempts) being

$$\Pr[Y_j = m | L] = \exp(-Lg) \frac{(Lg)^{m-1}}{(m-1)!}. \tag{11}$$

According to the extreme value theory [11], the maximum Y_j defined as $Y_{\max} \triangleq \max\{Y_j\}_{j=1}^{\infty}$ belongs to the maximum domain of attraction (MDA) of one of the three possible distribution types; namely the Gumbel, Fréchet or Weibull. According to [12, Thm. 3], the Poisson distribution is recoverable[3] in the MDA of Gumbel distribution yet not uniquely recovered. This occurs due to the discrete nature of Poisson samples. Nevertheless, for a sufficiently high parameter, say $Lg > 10$ in the parent PMF of Y_j as per (11), Poisson distribution resembles the Gaussian distribution. As shown in the numerical results section, the assumption of $Lg > 10$ is indeed a reasonable assumption since it reflects parametric values of practical interest. Insightfully, the latter distribution belongs to the MDA of Gumbel distribution and is uniquely recovered therein [12, Thm. 5]. Hence, $Y_{\max}$ converges to the Gumbel distribution.

The Gumbel distribution function is defined as

$$F_{\mathrm{G}}(x; \mu_{\mathrm{G}}, \sigma_{\mathrm{G}}) = \exp\left(-\exp\left(-\frac{x - \mu_{\mathrm{G}}}{\sigma_{\mathrm{G}}}\right)\right), \tag{12}$$

where μ_{G} and σ_{G} are the so-called location and scale parameter, respectively. Also, the expected value and variance of Gumbel distribution are given as $\mu_{\mathrm{G}} + \sigma_{\mathrm{G}}\epsilon$ and $\sigma_{\mathrm{G}}^2 \pi^2 / 6$, correspondingly. Via the method of moments-matching estimation, we obtain their respective estimates as

$$\hat{\mu}_{\mathrm{G}} = \overline{\Delta} - \epsilon \frac{\sqrt{6\mathcal{V}_{\Delta}}}{\pi}, \tag{13}$$

and

$$\hat{\sigma}_{\mathrm{G}} = \frac{\sqrt{6\mathcal{V}_{\Delta}}}{\pi}. \tag{14}$$

Thus, the distribution of the extreme (i.e., maximum) AoI, $\Delta_{\max}$, is approached by

$$F_{\Delta_{\max}}(x) \approx F_{\mathrm{G}}(x; \hat{\mu}_{\mathrm{G}}, \hat{\sigma}_{\mathrm{G}}). \tag{15}$$

Provided with the tail statistics of AoI, $\Delta_{\max}$ can be further quantified given a specific *level of risk*. By introducing the notion of confidence level a, we are able to characterize the level of risk in terms of AoI; for instance, a confidence level of a means that there is an $a\%$ possibility

---

[3]The explicit definition of a recoverable distribution is provided in [12, Eq. (1.1)].

that the worst-case AoI will not exceed a certain time threshold. Typical percentile values of the confidence level are $a = 99\%$ or $a = 95\%$. VaR is regarded as a key performance metric to quantify the effect of rare events. It is defined as [13, Def. 3.3]

$$\text{VaR}_{\Delta_{\max}}(a) \triangleq Q_{\Delta_{\max}}(a)$$

$$\approx \hat{\mu}_{\text{G}} - \hat{\sigma}_{\text{G}} \ln\left(-\ln(a)\right), \quad 0 < a < 1, \tag{16}$$

where $Q_{\Delta_{\max}}(\cdot)$ is the quantile function of $\Delta_{\max}$. However, VaR is an incoherent risk measure.[4] An alternative measure, which is indeed coherent and more reliable than VaR is the so-called CVaR. It is presented as [14, Prop. 15]

$$\text{CVaR}_{\Delta_{\max}}(a) \triangleq \frac{1}{1-a} \int_{a}^{1} \text{VaR}_{\Delta_{\max}}(a) dy$$

$$\approx \hat{\mu}_{\text{G}} + \frac{\hat{\sigma}_{\text{G}}}{1-a} \left[\text{li}(a) - a\ln\left(-\ln(a)\right) + \epsilon\right]. \tag{17}$$

It is noteworthy that $\text{CVaR}_{\Delta_{\max}}(a)$ is the expectation on the worst $(1-a)\%$ values of $\Delta_{\max}$, which reflects that it is more sensitive than VaR to the shape of the AoI distribution in the right tail. Notably, the term in brackets within (17) can be *a priori* computed and therefore is regarded as an offline operation.

IV. NUMERICAL RESULTS AND DISCUSSION

The derived analytical results are verified via numerical validation, whereas they are cross-compared with corresponding Monte-Carlo simulations. Without loss of generality and for ease of presentation, $d_e = d_c \triangleq d$ is considered. Also, the RF-to-DC conversion efficiency is set to be $\eta = 0.5$ while $w = 10^3$.

In Fig. 2, $g = 10$ and $g = 100$ correspond to a capacitor size $S = 0.2 \times 10^{-3}$, path-loss exponent $\alpha = 2.4$, transmit power $P_{\text{t}} = 40\text{dBm}$ as well as $d = 10\text{m}$. and $d = 26\text{m}$., respectively. Simulation data are cross-compared with the considered Gumbel distribution. Obviously, there is an almost perfect fit at extremely rare events in the right tail when maximum AoI is realized.

In Fig. 3, the performance of average AoI, VaR and CVaR is illustrated against different confidence levels. For all the considered cases, CVaR is higher than VaR since the former metric can capture the scenario of extreme AoI more rigorously than the latter one. Both the said

---

[4]A coherent risk metric satisfies certain properties regarding its supporting function; namely, translational invariance, sub-additivity, monotonicity, and homogeneity [13].
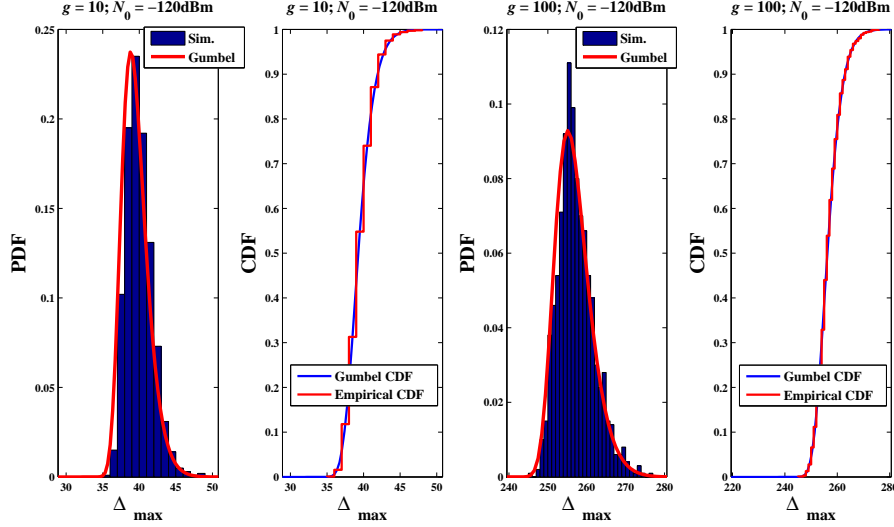
Fig. 2. Simulation data (entitled as 'Sim.' in legend) of the maximum AoI vs. the considered Gumbel distribution as per (15) for different system settings and a single-antenna transceiver with a target data rate $R = 0.1$ bps/Hz. PDF and CDF denote the probability density function and cumulative distribution function, respectively.

metrics maintain a notable difference from the average AoI (with respect to the number of required timeslots), even in the low confidence percentile of 70%. In fact, the latter difference becomes even more pronounced as the confidence level increases; this implies that for an almost sure convergence (i.e., for $a = 99\%$), taming quite a low AoI becomes a challenging task. Even for a more relaxed confidence level (e.g., when $a = 85\%$ or $a = 90\%$), there is a remarkable deviation between the most likely AoI (say, the average AoI) and its worst case counterpart. Moreover, the case when the capacitor size is $S = 10^{-4}$ reflects on a superior performance in terms of AoI in contrast to the case when $S = 10^{-3}$. It seems that (for this particular system setting) is better to operate using a lower transmission power profile at the source node so as to achieve a faster capacitor charge, which in turn reduces the total AoI. However, this holds true due to the moderately high received channel gains (i.e., a short link distance, $d = 10$m.) implying that the outage probability p is rather low.

Evidently, a promising solution to the aforementioned challenge (whenever the received channel gains are severely reduced) is to provide multiple antennas at the receiver side, as illustrated in Fig. 4. In order to obtain such an insight more emphatically, a massive antenna scale is applied. As expected, CVaR is reduced for the multiple antenna case since the received channel gains are being enhanced for higher N values, which in turn reflects on less packet errors and thus
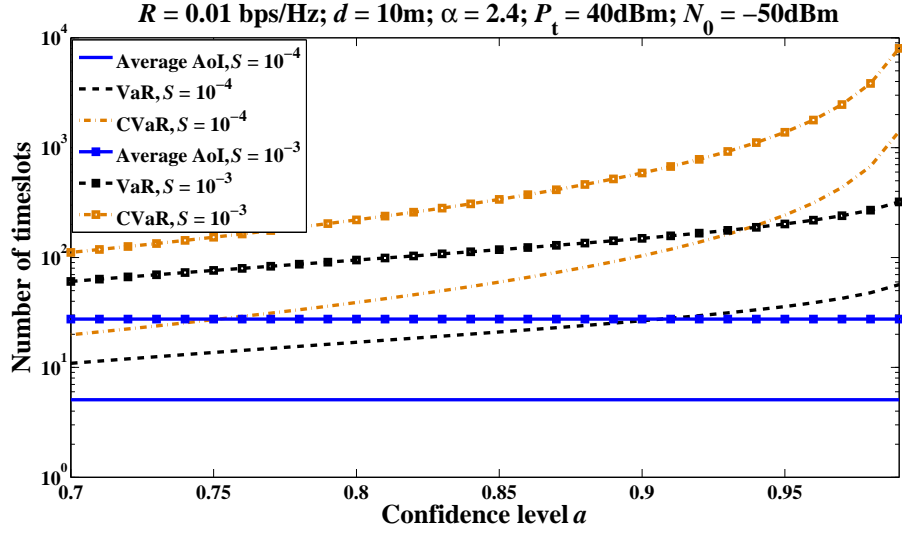
Fig. 3. Performance comparison of the average AoI, VaR and CVaR vs. various confidence levels for different system settings and a single-antenna transceiver.
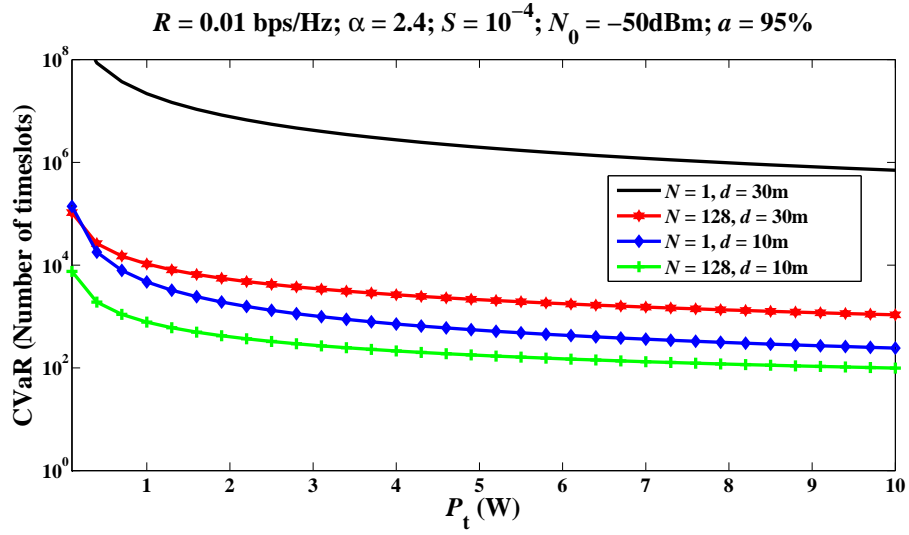

Fig. 4. Performance of CVaR vs. various transmission power levels (of the energy transmitter) for different system settings.

a shorter AoI. As demonstrated in Fig. 4, the performance difference is more impactful when channel fading and signal propagation attenuation is more severe (e.g., far-distant communication links).

V. CONCLUSIONS

The extreme AoI of a WPT communication system was analytically studied under independent Rayleigh channel fading conditions. It turns out that the extreme AoI of the considered system is suitably approached by the Gumbel distribution, while its associated parameters depend on the expected value and variance of AoI which were derived in simple closed-form expressions. Moreover, the VaR and CVaR metrics were obtained which demonstrate the deviation of the extreme AoI from its corresponding average counterpart. For a system setting of practical interest, it is observed that the capacitor size directly impacts on the extreme AoI, while an increasing number of receive antennas is beneficial only for severely degraded received channel gains (e.g., far-distant channel links).

APPENDIX

According to (8), the moments-function of Y should be computed. Recall that $Y = \sum_i X_i$ and X_i is a mutually independent RV $\forall i$. The probability mass function (PMF) of X_i defines the number of consecutive timeslots required for a fully-charged capacitor and is expressed as [3, Eq. (14)][5]

$$\Pr[X = k] = \exp(-g)\frac{g^{k-1}}{(k-1)!}. \tag{A.1}$$

Hence, the moment-function of X is given by

$$\mathbb{E}[X^n] \triangleq \sum_{k=1}^{\infty} k^n \Pr[X = k], \tag{A.2}$$

and it is specialized for the first four moments, respectively, as follows

$$\mathbb{E}[X] = 1 + g,$$
$$\mathbb{E}[X^2] = 1 + 3g + g^2,$$
$$\mathbb{E}[X^3] = 1 + 7g + 6g^2 + g^3,$$
$$\mathbb{E}[X^4] = 1 + 15g + 25g^2 + 10g^3 + g^4. \tag{A.3}$$

In addition, the moments-function of Y reads as

$$\mathbb{E}[Y^n] \triangleq \mathbb{E}\left[\left(\sum_{i=1}^{L} X_i\right)^n\right]$$

---

[5]Since X_i RVs are statistically independent, the ith index is dropped hereinafter.

$$= \mathbb{E}\left[\sum_{l_1+l_2+\cdots+l_L=n} \frac{n!}{l_1! l_2! \cdots l_L!} \prod_{i=1}^{L} X_i^{l_i}\right]. \tag{A.4}$$

Expanding the latter general expression for the first four moments and after some tedious yet straightforward manipulations, we arrive at the conditional moments of Y, namely

$$\mathbb{E}[Y|k] = \mathbb{E}\left[\left.\sum_{i=1}^{k} X_i \right| k\right] = k\mathbb{E}[X],$$

$$\mathbb{E}[Y^2|k] = \mathbb{E}\left[\left.\left(\sum_{i=1}^{k} X_i\right)^2 \right| k\right] = k\mathbb{E}[X^2] + k(k-1)\mathbb{E}^2[X],$$

$$\mathbb{E}[Y^3|k] = \mathbb{E}\left[\left.\left(\sum_{i=1}^{k} X_i\right)^3 \right| k\right]$$

$$= k\mathbb{E}[X^3] + 3k(k-1)\mathbb{E}[X^2]\mathbb{E}[X]$$

$$+ k(k-1)(k-2)\mathbb{E}^3[X],$$

$$\mathbb{E}[Y^4|k] = \mathbb{E}\left[\left.\left(\sum_{i=1}^{k} X_i\right)^4 \right| k\right]$$

$$= k\mathbb{E}[X^4] + 4k(k-1)\mathbb{E}[X^3]\mathbb{E}[X]$$

$$+ 6k(k-1)(k-2)\mathbb{E}[X^2]\mathbb{E}^2[X] + 3k(k-1)\mathbb{E}^2[X^2]$$

$$+ k(k-1)(k-2)(k-3)\mathbb{E}^4[X]. \tag{A.5}$$

The corresponding unconditional moments-function of Y in its general form is presented as

$$\mathbb{E}[Y^n] \triangleq \sum_{k=1}^{\infty} \mathbb{E}[Y^n|k]p^{k-1}(1-p), \tag{A.6}$$

which reflects on the case when $k-1$ unsuccessful attempts occurred before the kth successful transmission. By closely observing (A.5) and utilizing (A.6), calculations of the following type arise as

$$\sum_{k=1}^{\infty} kp^{k-1}(1-p) = (1-p)^{-1},$$

$$\sum_{k=1}^{\infty} k(k-1)p^{k-1}(1-p) = 2p(1-p)^{-2},$$

$$\sum_{k=1}^{\infty} k(k-1)(k-2)p^{k-1}(1-p) = 6p^2(1-p)^{-3},$$

$$\sum_{k=1}^{\infty} k(k-1)(k-2)(k-3)p^{k-1}(1-p) = 24p^3(1-p)^{-4}. \tag{A.7}$$

Finally, putting altogether the derivations of (A.3), (A.5) and (A.7) in (8), we reach at (9).

REFERENCES

[1] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, 2021.

[2] R. D. Yates, "The age of information in networks: Moments, distributions, and sampling," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5712–5728, 2020.

[3] I. Krikidis, "Average age of information in wireless powered sensor networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 628–631, 2019.

[4] O. M. Sleem, S. Leng, and A. Yener, "Age of information minimization in wireless powered stochastic energy harvesting networks," in *Proc. 54th Ann. Conf. Inf. Sci. Syst. (CISS)*, 2020, pp. 1–6.

[5] X. Jia, S. Cao, and M. Xie, "Age of information of dual-sensor information update system with HARQ chase combining and energy harvesting diversity," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2027–2031, 2021.

[6] C.-F. Liu and M. Bennis, "Taming the tail of maximal information age in wireless industrial networks," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2442–2446, 2019.

[7] M. K. Abdel-Aziz, S. Samarakoon, C.-F. Liu, M. Bennis, and W. Saad, "Optimized age of information tail for ultra-reliable low-latency communications in vehicular networks," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1911–1924, 2020.

[8] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*.   Academic Press, 2007.

[9] N. I. Miridakis, T. A. Tsiftsis, and G. C. Alexandropoulos, "MIMO underlay cognitive radio: Optimized power allocation, effective number of transmit antennas and harvest-transmit tradeoff," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 4, pp. 1101–1114, 2018.

[10] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*, 1st ed.   New York, NY, USA: Cambridge University Press, 2010.

[11] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*.   Springer Series in Operational Research and Financial Engineering, 2006.

[12] T. Shimura, "Discretization of distributions in the maximum domain of attraction," *Extremes*, vol. 15, pp. 299–317, 2012.

[13] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, pp. 203–228, 1999.

[14] M. Norton, V. Khokhlov, and S. Uryasev, "Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation," *Annals of Operations Research*, vol. 299, pp. 1281–1315, 2021.