APPENDIX

A. Protocol 1: Starting a new transmission

The following is the pseudo code of the protocol for initiating a new transmission, where DAT is the flag for data transmission requests, DSF is the data sending flag, t is the beginning of the next slot, and k is the next hopping channel in the single rendezvous coordination scheme or the hopping channel for the receiver in the multiple rendezvous coordination scheme.

```
Register initiation: DAT:=0, DSF:=0;
if a new data packet needs to be transmitted
   DAT := 1;
end if
if DAT=1
   predicting Pr(N_k(t) = 0), Pr(t_{k,off} > η|N_k(t) = 0);
end if
if Pr(N_k(t) = 0) ≥ τ_H AND Pr(t_{k,off} > η|N_k(t) = 0) ≥ θ
   generating a pseudo-random sequence;
else wait for the next time slot;
end if
if no RTS is heard before the corresponding mini slot
   sending RTS;
else wait for the next time slot;
end if
upon receiving CTS then
   DSF := 1;
if DSF=1
   DSF := 0;
   transmitting a data frame;
   DAT := 0 when transmission ends;
end if
```

B. Protocol 2: Spectrum handoff during a transmission

The following is the pseudo code of the protocol for the proactive spectrum handoff, where CSW is the channel switching flag, NUC and LSC are the number and the list of the candidate channels for data transmissions, respectively, and channel i is the current channel. As similar in Protocol 1, DAT is the flag for data transmission requests and DSF is the data-sending flag.

```
Register initiation: CSW:=0, DSF:=0, NUC:=0, LSC:=∅;
for j := 0, j ≤ M do
predicting Pr(N_j(t) = 0), Pr(t_{j,off} > η|N_j(t) = 0);
end for
if Pr(N_i(t) = 0) < τ_L AND DAT=1
   CSW := 1;
end if
if CSW=1
   for k := 0, k ≤ M do
      if Pr(N_k(t) = 0) ≥ τ_H AND Pr(t_{k,off} > η|N_k(t) = 0) ≥ θ
         NUC := NUC+1;
         LSC(NUC) := k;
      end if
   end for
end if
if LSC=∅
   wait for the next time slot;
elseif LSC ≠ ∅
   generating a pseudo-random sequence;
   broadcast channel availability information;
end if
upon receiving channel availability information then
   switching to the selected channel;
   starting the scanning radio;
if channel is busy
   wait for the next time slot;
else DSF := 1;
     CSW := 0;
end if
if DSF=1
   DSF := 0;
   transmitting a data frame;
   DAT := 0 when transmission ends;
end if
```

C. Derivation of the Spectrum Handoff Criteria for Biased-Geometric Traffic

According to Fig. 3 and (8), we calculate the spectrum handoff criteria proposed in Section III. We denote the finishing moment of the last PU packet as n_0 and the future time as slot n. Hence, the probability that channel i is idle and no PU arrival occurs between slot n_0+1 and n is given by

$$P_0 = 1 - \sum_{i=1}^{n-n_0} \lambda_n (1-\lambda_n)^{(i-1)}, \tag{11}$$

where λ_n is the normalized arrival rate. As shown in Fig. 3(b), the probability that channel i is idle and only one PU packet arrives between slot n_0+1 and n is

$$P_1 = \sum_{m=1}^{n-n_0-L} \left[1 - \sum_{i=1}^{n-n_0-m-L+1} \lambda_n(1-\lambda_n)^{(i-1)}\right] \lambda_n(1-\lambda_n)^{(m-1)}, \tag{12}$$

where m is the time slot at which a PU transmission starts and L is the length of a PU packet. Similarly, in Fig. 3(c), the probability that channel i is idle and h PU packets arrive between slot n_0+1 and n is

$$P_h = \sum_{m_h=h}^{n-n_0-hL} \left[1 - \sum_{i=1}^{n-n_0-m_h-hL+1} \lambda_n(1-\lambda_n)^{(i-1)}\right] \lambda_n^h(1-\lambda_n)^{(m_h-h)}. \tag{13}$$

Therefore, the total probability that channel i is idle at slot n is obtained as follows:

$$\Pr(N_i(n) = 0) = \sum_{i=0}^{U} P_i. \tag{14}$$

Secondly, due to the memoryless property of the geometric distribution, the probability that the duration of the idleness is longer than η slots on channel i is given by

$$\Pr(t_{i,off} > \eta | N_i(n) = 0) = 1 - \sum_{i=1}^{\eta} \lambda_n (1 - \lambda_n)^{(i-1)}. \quad (15)$$

D. Derivation of the Spectrum Handoff Criteria for Pareto Traffic

We follow the exact derivation procedure in Appendix *C* to calculate the spectrum handoff criteria of Pareto traffic. It is noted in (4) and (5) that the key is to obtain the expression of the distribution of the sum of W Pareto random variables (i.e., $V = \sum_{i=1}^{W} X_i$). In [43], the authors proved that, when $a = 1$, $0 < b < 2$ and $b \neq 1$, the CDF of V is given by

$$\Pr(\sum_{i=1}^{W} X_i > x) = \frac{-1}{\pi} \sum_{j=1}^{W} \binom{W}{j} (-\Gamma(1-b))^j \sin(\pi b j) \times \sum_{m=0}^{\infty} \frac{C_{W-j,m} \Gamma(m + bj)}{x^{(m+bj)}}, \quad (16)$$

where $\Gamma(\cdot)$ is the Gamma function and $C_{W-j,m}$ is the m-th coefficient in the series expansion of the $(W-j)$-th power of the confluent hyper-geometric function.

Therefore, the probability that channel i is idle and the probability that the duration of the idleness is longer than a frame size can be obtained by (4) and (5) if a is normalized to one and b is carefully selected.