Cognitive Coexistence between Infrastructure and Ad-hoc Systems

Stefan Geirhofer, Student Member, IEEE, Lang Tong, Fellow, IEEE, and Brian M. Sadler, Fellow, IEEE

Abstract—The rapid proliferation of wireless systems makes interference management more and more important. This paper presents a novel cognitive coexistence framework, which enables an infrastructure system to reduce interference to ad-hoc or peer-to-peer communication links in close proximity. Motivated by the superior resources of the infrastructure system, we study how its centralized resource allocation can accommodate the ad-hoc links based on sensing and predicting their interference patterns.

Based on an ON/OFF continuous-time Markov chain model, the optimal allocation of power and transmission time is formulated as a convex optimization problem and closed-form solutions are derived. The optimal scheduling is extended to the case where the infrastructure channel is random and rate constraints need only be met in the long-term average. Finally, the multi-terminal case is addressed and the problem of optimal sub-channel allocation discussed. Numerical performance analysis illustrates that utilizing the superior flexibility of the infrastructure links can effectively mitigate interference.

Index Terms—Cognitive Radio; Resource Allocation and Interference Management; Standards Coexistence; Dynamic Spectrum Access;

I. INTRODUCTION

The rapid growth of wireless networks makes interference an important performance impediment and motivates a careful study of coexistence. In unlicensed bands, where there is a lack of coordination among heterogeneous technologies, traditional forms of coexistence are based on statically separating systems, for example by having them operate in disjoint frequency bands. However, such static approaches are inefficient and not able to accommodate the projected growth in deployments.

The high interference conditions which are prevalent in unlicensed bands are in contrast to a fairly low average utilization, which results from the burstiness of traffic that is supported by these systems and the random medium access behavior typically employed. This contrast motivates dynamic approaches to interference management which can leverage unused spectrum opportunities.

Cognitive radio presents a new framework for analyzing this problem. Reconfigurable radio platforms enable a dynamic adaptation of transmission parameters based on monitoring the radio environment. In this way it is possible to reduce interference by detecting and predicting temporal activity patterns and avoiding transmission overlaps. Ultimately this may lead to a new cognitive coexistence paradigm.

This paper addresses a special case of cognitive coexistence involving two different types of networks: an infrastructure (IS) wide area network that shares spectrum with local, ad-hoc (AH) or peer-to-peer systems. Motivated by the superior communication resources of the IS system, we analyze how its flexible, centralized resource allocation can accommodate the AH links based on sensing and predicting their interference patterns. Despite adapting its resource allocation based on sensing results, the IS system minimizes interference to the AH network subject to maintaining a specified quality-of-service (QoS) level for its users.

This approach is different from typical dynamic spectrum access (DSA) problem formulations in which a secondary system exploits spectrum opportunities left over by a primary system, subject to the constraint that no significant interference is created (see [1] for a review). While both approaches are hierarchical, spectrum property rights in DSA mandate that secondary users adapt to spectrum licensees. In contrast, in cognitive coexistence it is possible to use the flexibility of the primary system to accommodate a secondary (perhaps lower priority) system while maintaining a desired primary system performance level. This could be viewed as a “best-effort” approach toward interference management and coexistence.

This framework is relevant to a number of practical setups. For example, the convergence of wide and local area networks has received increasing interest. The coexistence of IEEE 802.16 and IEEE 802.11 systems in unlicensed bands is such an example of practical importance [2], [3]. Similarly, incorporating peer-to-peer connectivity into cellular networks [4] or accommodating femto-cell base stations [5] are potential scenarios where this work could be applied. Yet other applications arise in the military domain, where the coexistence of high and low priority links is a fundamental concern [6].

A. Main contribution

This paper addresses the coexistence of two wireless networks that operate on different spatial scales: a longer range IS network that interferes with local AH links. We address the question of whether the IS network can accommodate surrounding AH links by allocating power and transmission time judiciously. Specifically, we make the following contributions:
The problem of optimal power and transmission time allocation is formulated as a convex program and the optimal frame-level solution is derived. Based on optimality conditions, a solution algorithm with guaranteed convergence and low complexity is introduced. The structure of the optimal solution is studied to provide further intuition.

- The frame-level problem is relaxed to the case of average rate constraints, in which statistical knowledge of the activity patterns of the AH users and the IS channel coefficients are used to allocate resources in both frequency and time.
- Finally, a scenario with multiple IS users is considered in which the IS base station allocates sub-channels based on average interference metrics and users perform optimal power and transmission time allocation based on sensing results. A comparison with conventional sub-channel allocation methods shows that heuristics may yield performance close to optimal.

These contributions are corroborated by numerical results which demonstrate that judicious transmission time allocation can mitigate interference effectively.

B. Related work

Optimal resource management in multiuser multicarrier wireless systems has been well-studied for both downlink and uplink cases; see [7], [8], [9] for an overview of the topic. In cognitive radio networks, optimal resource allocation is more challenging because it needs to incorporate interference constraints, which protect the primary system from harmful interference. Typical formulations aim at finding a power and sub-channel allocation which maximizes the throughput of the cognitive radio system while meeting interference and power constraints. Recent work in this area includes [10], [11]. In addition to meeting interference constraints, spectrum sharing and self-coexistence within the cognitive radio network also need to be addressed. Contributions in this area include [12], [13], [14], [15]. Within the framework of cognitive coexistence, the optimal power allocation based on knowledge of the interference channel has been addressed in [16] by the authors of this paper. Previous work on improving the coexistence among local and personal area networks includes [17], [18] in which a cognitive frequency hopping protocol is derived based on temporal activity models. To the best of the authors’ knowledge, interference-aware power and transmission time allocation based on predicting temporal activity patterns has not been addressed before.
We model the time behavior of each AH band by a two-state ON/OFF continuous time Markov chain (CTMC). The holding times in both ON and OFF state are exponentially distributed with parameters $\mu$ for the ON state and $\lambda$ for the OFF state. Therefore, if an AH link is detected to be in a certain state at time $t_0$, then its transition matrix for time $t_0 + \tau$ is given by

$$
P(\tau) = \frac{1}{\lambda + \mu} \begin{bmatrix} \mu + \lambda e^{-(\lambda+\mu)\tau} & \lambda - \lambda e^{-(\lambda+\mu)\tau} \\ \mu - \mu e^{-(\lambda+\mu)\tau} & \lambda + \mu e^{-(\lambda+\mu)\tau} \end{bmatrix}, \tag{1}
$$

which follows directly from the definition of a CTMC \[19\], p.391. Therefore, the probability of an AH link being ON at time $t_0 + \tau$, conditioned on having it observed in the ON (OFF) state at time $t_0$ is given by the lower right (upper right) entry in the matrix above.

Modeling AH links based on a two-state CTMC approximates the carrier sense random medium access typically employed in such systems. This modeling approach has been used in related publications \[18, 20\] and solidified by a measurement-based analysis of WLAN traffic \[17\]. The CTMC assumption strikes a good tradeoff between model accuracy and the analytical tractability that is needed in the subsequent sections.

**Infrastructure system:** The IS system operates in the same frequency band as the AH network and evolves in frames of fixed duration $T$. At the beginning of each frame, spectrum sensing is used to detect the ON/OFF activity of the AH bands, and based on the sensing outcome, power and transmission time are assigned; see Fig. \[13\]

Throughout the paper we assume perfect sensing, that is, sensing outcomes are always accurate and the overhead associated with sensing is negligible\[6\]. Due to the proximity of IS and AH systems, the detection task is conceptually similar to the carrier sense employed in systems such as IEEE 802.11.

Based on the sensing result at the beginning of each frame, the IS system allocates power and transmission time on a sub-channel basis. This is conceptually similar to the allocation of time/frequency resource blocks in broadband cellular systems based on OFDMA. The case where a subset or even all sub-channels need to share the same timing allocation (for example when transmissions in the entire band can only be turned on or off) has worse performance in general. Nevertheless, a similar solution approach remains applicable.

The IS system minimizes interference subject to maintaining rate requirements for its client. The rate that is supported by a specific sub-channel is modeled based on a channel capacity formulation,

$$
\sum_n \rho_n \log \left( 1 + \kappa \frac{p_n|h_n|^2}{\rho_n N_0} \right) = \sum_n \rho_n \log \left( 1 + \frac{p_n \beta_n}{\rho_n} \right), \tag{2}
$$

where $p = [p_1, \ldots, p_N]^T$ denotes the power allocation, $\rho = [\rho_1, \ldots, \rho_N]^T$ represents the transmission time allocation.

$N_0$ is the noise power, $\kappa$ a normalization factor, and $\beta_n$ is introduced for notational convenience\[4\].

### B. Interference metrics and scheduling assumptions

The interference between IS and AH networks is modeled by the average temporal overlap between both systems. Based on the sensing result at the beginning of the frame and knowledge of the CTMC parameters of the AH links, transmission time and power are allocated.

The allocation of transmission time consists of specifying duration and placement of the transmission within the current frame. We first show that it is optimal to transmit at the beginning (the end) of the frame if the sensing outcome is idle (busy).

**Lemma 1:** Assume that a $\rho$ fraction of transmission time needs to be allocated to a sub-channel, on which the AH user’s ON/OFF behavior is modeled by the CTMC \[1\]. Based on a sensing outcome at the beginning of the frame, the minimum expected overlap with the ON period of the AH user is achieved by

- transmitting at the beginning of the frame (i.e., during $[0, \rho T]$) if the sensing outcome was idle and
- transmitting at the end of the frame (i.e., during $[(1 - \rho)T, T]$) if the sensing outcome was busy.

**Proof:** see appendix.

Based on Lemma 1, we derive the expected time overlap between IS and AH transmissions, conditioned on the sensing result $y \in \{0,1\}$ at the beginning of the frame. Consider a sub-channel $n$, which overlaps with AH band $i = g(n)$. Then, the activity of AH user $i$ is given by the CTMC $\{X_i(\xi), \xi \geq 0\}$ with parameters $\lambda_i$ and $\mu_i$. Transmitting for a $\rho$ fraction of the frame, leads to the expected time overlap

$$
\phi_{n,0}(\rho) = \frac{1}{T} \mathbb{E} \left\{ \int_0^{\rho T} \mathbb{I}_{\{X(\xi)=1\}} d\xi \middle| X(0) = 0 \right\} = \frac{1}{T} \int_0^{\rho T} \Pr(X(\xi) = 1|X(0) = 0) d\xi \tag{3}
$$

if the sensing result was idle. By substituting \[1\] it is then easy to show that

$$
\phi_{n,0}(\rho) = \frac{\lambda_i}{(\lambda_i + \mu_i)T} \left( \rho T + \frac{1}{\lambda_i + \mu_i} e^{-(\lambda_i + \mu_i)\rho T} - 1 \right). \tag{4}
$$

In the case of a busy sensing result we obtain

$$
\phi_{n,1}(\rho) = \frac{\lambda_i}{(\lambda_i + \mu_i)T} \left( \rho T + \frac{\mu_i}{\lambda_i + \mu_i} e^{-(\lambda_i + \mu_i)\rho T} (e^{(\lambda_i + \mu_i)\rho T} - 1) \right). \tag{5}
$$

This derivation made use of the fact that the sensing results of two IS sub-channels are either perfectly correlated (if they overlap with the same AH band) or statistically independent (if they overlap with different bands). Therefore, the prediction

\[2\] The above formulation encompasses a channel capacity formulation (for $\kappa = 1$) as well as the case of variable-rate M-QAM in which case $\kappa = 1.5/(-\ln \text{BER})$ is chosen such that a target BER is met \[21\].
performance of a specific sub-channel cannot be improved by using sensing results from other sub-channels.

**Lemma 2:** The functions \( \phi_{n,0}(\rho) \) and \( \phi_{n,1}(\rho) \) are strictly convex and increasing in \( \rho \).

**Proof:** Both \( \phi_{n,0}(\rho) \) and \( \phi_{n,1}(\rho) \) are nonnegative linear combinations of a convex and a strictly convex function. One is linear, the other an exponential function with nonzero exponent. The monotonicity can easily be verified by differentiation.

### III. Optimal Frame-Level Allocation

Consider a single IS client, which minimizes the time overlap between IS and AH transmissions subject to maintaining a rate constraint across the IS channel to the IS base station. Mathematically, this leads to problem \( \textbf{P1} \)

\[
\begin{align*}
\min_{\mathbf{p}, \mathbf{\rho}} & \quad \sum_{n} \phi_{n,y_n}(\rho_n) \\
\text{s.t.} & \quad \sum_{n} \rho_n \log \left(1 + \frac{p_n \beta_n}{\rho_n}\right) \geq R \\
& \quad \sum_{n} p_n \leq P \\
& \quad p_n \geq 0, \quad 1 \leq n \leq N \\
& \quad 0 \leq \rho_n \leq 1, \quad 1 \leq n \leq N,
\end{align*}
\]

with rate constraint (7) and power constraint (8). It is straightforward to show that \( \textbf{P1} \) is a convex optimization problem since the objective function is convex (by Lemma 2), the rate constraint (once rewritten in standard form) is convex by the perspective property [22], and all other constraints are linear.

A solution to \( \textbf{P1} \) can be found by general solution techniques in polynomial time [22]. For this specific problem, however, it is possible to show a special structure that enables us to gain further insight into the problem.

#### A. Optimality conditions and solution structure

The solution structure is obtained by introducing Lagrange multipliers \( \gamma \) and \( \epsilon \) for the rate and power constraint, respectively. This leads to the Lagrangian

\[
L(\mathbf{p}, \mathbf{\rho}; \gamma, \epsilon) = \sum_{n} \phi_{n,y_n}(\rho_n) + \gamma \left[ R - \sum_{n} \rho_n \log \left(1 + \frac{p_n \beta_n}{\rho_n}\right) \right] + \epsilon \left[ \sum_{n} p_n - P \right].
\]

The Karush-Kuhn-Tucker (KKT) optimality conditions are then given by the constraints (7)-(10) of \( \textbf{P1} \), nonnegativity constraints for the Lagrange multipliers, \( \gamma \geq 0, \epsilon \geq 0 \), the slackness conditions

\[
\begin{align*}
\gamma \left[ R - \sum_{n} \rho_n^* \log \left(1 + \frac{p_n^* \beta_n}{\rho_n^*}\right) \right] & = 0 \quad \text{(12)} \\
\epsilon \left[ \sum_{n} p_n^* - P \right] & = 0, \quad \text{(13)}
\end{align*}
\]

and

\[
\frac{\partial L(\mathbf{p}, \mathbf{\rho}; \gamma, \epsilon)}{\partial \rho_n} \bigg|_{\rho_n = \rho_n^*} \begin{cases} > 0, & \rho_n^* > 0 \\
= 0, & \rho_n^* = 0 \\
< 0, & \rho_n^* = 1
\end{cases} \quad \text{(14)}
\]

Expressions (14) and (15) can be understood on an intuitive level by noting that for \( \rho_n^* \) and \( \rho_n^* \) to be minimizers of \( L(\mathbf{p}, \mathbf{\rho}; \gamma, \epsilon) \), their partial derivative must equal zero unless they lie on the boundary of the feasible set.

**Solution structure for \( \rho_n^* \):** By substituting (11) into (14) and solving for \( \rho_n^* \) we arrive at

\[
\rho_n^* = \rho_n \left( \nu - \frac{1}{\beta_n} \right)^+, \quad \text{(16)}
\]

where \( \nu := \gamma / \epsilon \) has been introduced to simplify notation in what follows. For any fixed value of \( \rho_n \), (13) represents a water filling solution [23].

**Solution structure for \( p_n^* \):** The optimal transmission time allocation is obtained by substituting (11) and (16) into (15). For an idle sensing result, \( y_n = 0 \) we obtain,

\[
\begin{align*}
p_n^* & = \begin{cases} \frac{1}{\lambda_i + \mu_i} \log \frac{1}{1 + (\lambda_i + \mu_i)T}, & \gamma h_n(\nu) \leq \zeta_{0,i}, \\
1, & \text{o.w.}
\end{cases} \quad \text{(17)}
\end{align*}
\]

where \( \zeta_{0,i} = \lambda_i / (\lambda_i + \mu_i) (1 - \exp(-\lambda_i T)) \) and \( i = g(n) \) denotes the AH sub-band that overlaps with sub-channel \( n \). In the above equation we have defined

\[
\gamma h_n(\nu) := \left[ \log(\nu \beta_n) \right]^+ - \frac{(\nu \beta_n - 1)^+}{1 + (\nu \beta_n - 1)^+},
\]

to simplify notation. Similarly, we can obtain the solution structure for the case of a busy sensing result, \( y_n = 1 \),

\[
\begin{align*}
p_n^* & = \begin{cases} 0, & \gamma h_n(\nu) < \zeta_{1,i}, \\
1 + \frac{\log(\nu \beta_n)}{(\lambda_i + \mu_i)T}, & \zeta_{1,i} \leq \gamma h_n(\nu) \leq 1 \\
1, & \gamma h_n(\nu) > 1,
\end{cases} \quad \text{(19)}
\end{align*}
\]

where \( \zeta_{1,i} = \lambda_i / (\lambda_i + \mu_i) (1 + \mu_i / \lambda_i) \exp(-\lambda_i T) \). Note that these closed-form expressions depend on the Lagrange multipliers only through the term \( \gamma h_n(\nu) \) which does not depend on the AH activity parameters \( \lambda_i \) and \( \mu_i \). Further, the transmission time allocations (17)-(19) are monotonic with respect to this term.

#### B. Iterative solution algorithm for \( \gamma \) and \( \nu \)

To find the optimal power and transmission time allocation based on the above closed-form expressions, we present an algorithm for finding the pair \( [\gamma^*, \nu^*] \), which corresponds to the optimal solution of \( \textbf{P1} \).

For any pair \( [\gamma, \nu] \) the power allocation \( \mathbf{p}(\gamma, \nu) \) and transmission time allocation \( \mathbf{\rho}(\gamma, \nu) \) define the optimal solution to \( \textbf{P1} \) with modified rate constraint

\[
R(\gamma, \nu) := \sum_{n} \rho_n(\gamma, \nu) \left[ \log(\nu \beta_n) \right]^+, \quad \text{(20)}
\]

and modified power constraint

\[
P(\gamma, \nu) := \sum_{n} \rho_n(\gamma, \nu) \left( \nu - \frac{1}{\beta_n} \right)^+, \quad \text{(21)}
\]

where \( p_n(\gamma, \nu) \) is given by (17) or (19) (depending on the sensing result). The fact that this solution is optimal for rate constraint \( R(\gamma, \nu) \) and power constraint \( P(\gamma, \nu) \) follows directly from the KKT optimality conditions, which are necessary and sufficient for convex optimization problems [22].

Based on the above, finding the pair \([\gamma^*, \nu^*] \) corresponding to the given rate constraint \( R \) and power constraint \( P \) could theoretically be performed by searching all pairs \([\gamma, \nu] \) in this section we show, however, that \( R(\gamma, \nu) \) and \( P(\gamma, \nu) \) exhibit some monotonicity which enables us to use the bisection method for finding \([\gamma^*, \nu^*] \) with guaranteed convergence and low complexity. We first study the case of keeping \( \gamma \) fixed and adjusting \( \nu \) such that the rate constraint is met with equality. Then, we show that the allocated sum power decreases with \( \gamma \).

**Adjusting \( \nu \) to meet the rate constraint:** We first consider the case of adjusting \( \nu \) such that \( R(\gamma, \nu) = R \) while keeping \( \gamma \) fixed. It is easy to verify that for any \( n, h_n(\nu) \) is nondecreasing in \( \nu \). Therefore, \( p_n(\gamma, \nu) \) increases with \( \gamma \) as well as can be seen from (17) and (19). Further, since \( p_n(\gamma, \nu) \) increases with \( \nu \) for fixed \( \gamma \), so does \( R(\gamma, \nu) \). We can exploit this property to find the \( \nu \) for which \( R(\gamma, \nu) = R \) by the bisection method. First, we can find upper and lower bounds, \( \nu_u \) and \( \nu_l \), for this value. These bounds are guaranteed to exist since \( R(\gamma, \nu) \rightarrow \infty \) for \( \nu \rightarrow \infty \) and \( R(\gamma, \nu) \rightarrow 0 \) for \( \nu \rightarrow 0 \). Once these bounds have been obtained the bisection method iteratively finds \( \nu^* \) with guaranteed convergence.

**Adjusting \( \gamma \) to meet the power constraint:** Having obtained an algorithm for finding \( \nu \) for arbitrary \( \gamma \) such that the rate constraint is satisfied, we study the behavior of the power constraint as \( \gamma \) is adjusted. As \( \gamma \) is varied, we continue to adjust \( \nu \) such that the rate constraint is satisfied at all times. The pair of Lagrange multipliers is therefore given by \([\gamma, \nu^*(\gamma)]\).

The slackness conditions imply that at the optimal solution both rate and power constraints are met with equality. From (20) we observe that decreasing \( \gamma \) requires increasing \( \nu \) in order to continue meeting the rate constraint. Further, decreasing \( \gamma \) reduces the objective function because \( \log(\hat{\nu}\beta_n) \geq \log(\nu\beta_n) \) for \( \hat{\nu} \geq \nu \), enables us to reduce \( \rho_n(\gamma, \nu^*(\gamma)) \) for at least some \( n \).

Since decreasing \( \gamma \) requires increasing \( \nu \), the allocated sum power increases as \( \gamma \) decreases. Intuitively, a constant rate constraint (20) requires (21) to increase because the term \([\nu^*(\gamma) - \frac{1}{2}r^2] \) increases faster than \([\log(\nu^*(\gamma)\beta_n)]^+ \). The following lemma shows this rigorously.

**Lemma 3:** The sum power \( \sum_n p_n \) associated with allocation \([\gamma, \nu^*(\gamma)]\) is a decreasing function in \( \gamma \).

**Proof:** see appendix.

Lemma 3 enables us to find \( \gamma^* \) again by the bisection method. Assuming that \( \textbf{P1} \) is feasible which we will assume hereafter, there exist bounds \( \gamma_u \) and \( \gamma_l \) such that \( P(\gamma_u, \nu^*(\gamma_u)) \leq P \leq P(\gamma_l, \nu^*(\gamma_l)) \). Therefore, by starting the bisection method from these points we can find the pair \([\gamma^*, \nu^*] \) with guaranteed convergence. The solution algorithm is shown in detail in Fig. 2. The inner loop (lines 4–14) correspond to finding \( \nu^*(\gamma) \), whereas the outer loop finds \( \gamma^* \).

**Algorithm 1: Solution Algorithm**

1. **Initialization.** Obtain bounds \( \nu_l, \nu_u, \gamma_l, \gamma_u \).
2. **repeat**
   3. \( \hat{\gamma} \leftarrow (\gamma_u - \gamma_l)/2 \);
   4. **repeat**
   5. \( \hat{\nu} \leftarrow (\nu_u - \nu_l)/2 \);
   6. Find time allocation \( \rho_n(\hat{\nu}) \) using (17);
   7. Find power allocation \( p_n(\hat{\gamma}, \hat{\nu}) \) using (16);
   8. Compute achievable rate \( r(p_n, \rho_n) \) using (2);
   9. if \( r(p_n, \rho_n) \geq R \) then
      10. \( \nu_l \leftarrow \hat{\nu} \)
   11. else
      12. \( \nu_l \leftarrow \hat{\nu} \)
   13. **end**
   14. until \( 0 \leq R - r(p_n, \rho_n) \leq \epsilon_R \);
   15. Find time allocation \( \rho_n(\hat{\gamma}, \nu) \) using (17);
   16. Find power allocation \( p_n(\hat{\gamma}, \nu) \) using (16);
   17. if \( \sum_n p_n \) then
      18. \( \hat{\gamma} \leftarrow \gamma \)
   19. else
      20. \( \gamma_u \leftarrow \hat{\gamma} \)
   21. **end**
   22. until \( 0 \leq P - \sum_n p_n \leq \epsilon_p \);

Fig. 2: Algorithm for finding the optimal Lagrange multipliers \( \gamma \) and \( \nu \) for problem (5)–(10). The inner loop (lines 4–14) find \( \nu^*(\gamma) \) which satisfies the rate constraint (7). The outer loop determines \( \gamma^* \), which satisfies the power constraint (8).

**C. Properties of optimal allocations**

Beyond simplifying solution algorithms, the structured solutions also enable us to make some qualitative statements about the optimal resource allocation.

**Ordering in terms of IS channel quality:** We first investigate the ordering with respect to the IS channel coefficients \( \beta_n \). This scenario is shown in Fig. 3a for \( N = 4 \) sub-channels with coefficients \( \beta = [0.9, 1.1, 1.5, 1.5]^T \). We observe that, for any rate constraint, more transmission time is allocated to the sub-channel with higher channel coefficient, i.e., \( \beta_i \geq \beta_j \Rightarrow \rho_i \geq \rho_j \) provided all other parameters and the sensing results are identical. On an intuitive level, this result captures the fact that in channels with high \( \beta \), we can achieve the same rate in a shorter transmission duration using the same amount of power. Mathematically, the result follows from the monotonicity of the optimal solution and \( h_n(\nu) \).

**Ordering in terms of sensing results:** Similar to the IS channel, the optimal transmission time allocation can be ordered with respect to the sensing results. If channels have the same IS channel coefficient, \( \beta_i = \beta_j \) but different sensing results then it is preferable to allocate more transmission time to the idle channel, i.e., \( \rho_i \geq \rho_j \). This is illustrated in Fig. 3b for \( N = 4, \beta = [0.9, 1.1, 1.1, 1.1]^T \) and \( y = [0.1, 0, 1, 1]^T \). It is also interesting to note that some transmission time is allocated to frames with busy sensing results even when idle frames are not yet used to the maximum extent.
IV. OPTIMAL AVERAGE RESOURCE ALLOCATION

Problem formulation P1 required that rate and power constraint are met in every frame, even if sensing outcome or IS channel quality are disadvantageous. In practical systems, satisfying rate constraints at the frame-level is usually unnecessary; it suffices to maintain average rate constraints across time. This less stringent requirement can be used to further reduce interference by allocating less transmission time during frames with adversarial channel/interference conditions, while compensating for the rate decrease during frames with better conditions. Ultimately, this leads to an improved resource allocation across both frequency (the sub-channels of the IS system) and time (consecutive frames of the IS system).

This section introduces such an average rate formulation by averaging across the temporal activity of the AH network and random IS channel coefficients. Further, this section introduces two reference schemes that help to put the performance of the optimal resource allocation in perspective.

A. Formulation and solution structure

The average rate formulation requires associating probabilities with all possible sensing outcomes. While there are a total of $N$ sub-channels available, the sensing outcomes for sub-channels that overlap with the same AH band will be identical. Therefore, for $M$ sub-bands, there are a total of $2^M$ possible sensing outcomes. Let the set of all possible sensing outcomes be represented by $\mathcal{Y} = \{0,1\}^M$ where $\mathcal{Y} = [y_1, \ldots, y_M]^T \in \mathcal{Y}$ denotes the sensing outcome per sub-band.

Problem P2 of optimally allocating power and transmission time then becomes

$$\min_{\rho_{n,y}} \sum_{\mathcal{Y}} \eta_{\mathcal{Y}} \sum_n \rho_{n,y(i)}(\rho_{n,y}) \quad (22)$$

subject to

$$\sum_{\mathcal{Y}} \eta_{\mathcal{Y}} \sum_n \rho_{n,y} \log \left(1 + \frac{p_{n,y} \beta_n}{\rho_{n,y}} \right) \geq R \quad (23)$$

$$\sum_{\mathcal{Y}} \eta_{\mathcal{Y}} \sum_n p_{n,y} \leq P \quad (24)$$

$$p_{n,y} \geq 0, \quad \forall \mathcal{Y}, 1 \leq n \leq N \quad (25)$$

$$0 \leq \rho_{n,y} \leq 1, \quad \forall \mathcal{Y}, 1 \leq n \leq N, \quad (26)$$

where $\eta_{i,0} = \mu_i/(\lambda_i + \mu_i)$, $\eta_{i,1} = \lambda_i/(\lambda_i + \mu_i)$, and due to the independence of the AH sub-bands, $\eta_{\mathcal{Y}} = \prod_{i=1}^M \eta_{i,y(i)}.$

Note that this optimization problem has $2^M$ as many decision variables because power and transmission time allocation may be different for every possible sensing outcome. The fact that the decision variables grow exponentially with $M$ is not of major concern, because $M$ (the number of parallel AH bands) is typically quite small (in the order of one to five).

Problem P2 can be solved similar to problem P1. In particular, by forming the Lagrangian, introducing Lagrange multipliers $\gamma$ and $\epsilon$, and taking the derivative with respect to the decision variables, we obtain a similar solution structure as in the frame-level problem.

B. Reference schemes

Having incorporated random sensing outcomes into our formulation, we introduce two reference schemes in order to put the performance of the optimal resource allocation in perspective.

No sensing case: As a first benchmark, consider an approach that allocates power but does not perform any transmission time optimization. This case corresponds to conventional resource management in IS systems, which simply allocates power to make the best use of the uplink channel. Mathematically, this is formulated as minimizing $\sum_n p_n$ subject to the constraints (17–19). We assume that for any sub-channel with $p_n > 0$, the sub-channel is used for the entire frame duration. Unused carriers for which $p_n = 0$ are not allocated any transmission time.

Fig. 3: Structure of the optimal transmission time allocation for varying rate constraint. The solution can be ordered with respect to IS channel coefficients or sensing outcomes.
Idle-frame allocation: Another possible reference scheme performs spectrum sensing but allocates resources in a suboptimal way. Specifically, consider allocating the entire frame by setting $\rho_n = 1$ for all idle sub-channels while completely avoiding busy sub-channels by setting $\rho_n = 0$ for all $n$ with $y_n = 1$. In the average rate formulation, this method can be formulated mathematically as minimizing $\sum_{\gamma \in \mathcal{Y}} \eta_{\gamma} \rho_n$ subject to (23)-(26) and the additional constraint that subchannels with busy sensing outcome are never allocated. Note that the above optimization problem may be infeasible even when $\mathbf{P2}$ is feasible because we are imposing the additional restriction of never transmitting during busy frames. To ensure that the reference scheme is always feasible when $\mathbf{P2}$ is, we force allocation to busy channels if the resulting optimization would otherwise be infeasible.

C. Allocation for random IS channels

The previous section considered average rate constraints with respect to temporal AH activity but fixed IS channel $\beta$. This section further extends the analysis to the case of random IS channel coefficients $\beta$. The optimization problem $\mathbf{P3}$ is given by

$$\min_{\mathbf{p}(y, \beta), \rho(y, \beta)} \int_\beta \eta_{\gamma} \sum_n \phi_{n, y_{\beta(n)}}(\rho_n(y, \beta)) dF(\beta)$$

subject to

$$\int_\beta \eta_{\gamma} \sum_n \rho_n(y, \beta) \log(1 + \frac{\rho_n(y, \beta)\beta_n}{\rho_n(y, \beta)}) dF(\beta) \geq R$$

$$\int_\beta \eta_{\gamma} \sum_n \rho_n(y, \beta) dF(\beta) \leq P$$

$$\rho_n(y, \beta) \geq 0, \quad \forall y \in \mathcal{Y}, 1 \leq n \leq N$$

$$0 \leq \rho_n(y, \beta) \leq 1, \quad \forall y \in \mathcal{Y}, 1 \leq n \leq N,$$

where the decision variables $\mathbf{p}(y, \beta)$ and $\rho(y, \beta)$ correspond to the power and transmission time allocation that is used for sensing outcome $y$ and IS channel condition $\beta$ and $F(\beta)$ is the cumulative distribution function of $\beta$. By again forming the Lagrangian and computing the derivative with respect to the decision variables, it is easy to show that the structured solutions (14) and (17)-(19) again hold. Therefore, it is again possible to express the allocation as a function of the Lagrange multipliers $[\gamma, \nu]$. The rate constraint can then be evaluated by

$$\int_\beta \eta_{\gamma} \sum_n \rho_n(y, \beta)[\log(\nu_{\beta(n)})]^+ dF(\beta)$$

and the allocated sum power is given by

$$\int_\beta \eta_{\gamma} \sum_n \rho_n(y, \beta) \left(\nu - \frac{1}{\rho_n}\right)^+ dF(\beta).$$

While the above integrals can only be evaluated numerically, it is possible to again find the optimal solution via the bisection method.

D. Numerical results

This section presents numerical performance results for the optimal average rate resource allocation and compares them to the reference schemes introduced in this section. The results were obtained for $N = 5$ sub-channels and a single AH subband $M = 1$. The prediction parameters were $\lambda = \mu = 1 \text{s}^{-1}$ and the IS channel coefficients were flat Rayleigh fading and statistically independent. We further assume a block fading scenario in which the IS channel varies slowly compared to the frame duration.

The performance for fixed IS channel and random sensing results is shown in Fig. 4a for $T = 1 \text{s}$ and in Fig. 4b for $T = 0.1 \text{s}$. The plot shows the average transmission overlap between IS and AH network versus the achieved IS rate (note that the achieved IS rate and not the rate constraint would otherwise be infeasible.)
is plotted. The performance results are averaged over 100 realizations of the IS channel. Since the IS channel is not modeled statistically, it is inevitable that for some realizations of $\beta$ problem $P2$ is infeasible. The outage probability, which is identical for all three schemes, is therefore shown in Fig. 4c to put the results in perspective. Typical outage probabilities of approximately 10% correspond to an IS rate of about 0.7 bps. At this rate, plots (a) and (b) show that a significant performance gain is achieved by performing sensing-based transmission time allocation.

The performance ordering reflects our expectations. The idle-frame allocation scheme outperforms the no-sensing case but shows a quite significant performance gap with respect to the optimal allocation, especially for low IS rates. Further, all curves show increasing interference as the IS rate increases. This is expected, since high IS rates prevent the IS system from being able to accommodate the AH links. The plots also show that idle-frame allocation and no-sensing scheme converge for high IS rates, because allocating only idle frames is almost always infeasible (and therefore busy frames typically need to be used as well).

By comparing Fig. 4a ($T = 1$) and Fig. 4b ($T = .1$), we observe that while the performance of the optimal scheme does not change significantly, the idle-frame reference performs much better. This is intuitive, because by reducing the frame length, it is easier to “fill up” the idle periods of the AH network. The performance of the no-sensing scheme remains unaltered and is the same in both figures.

The performance for average IS channel coefficients and random AH behavior is shown in Fig. 5 which compares the solution of $P3$ with the same reference schemes. We can observe that by exploiting the channel variability and allocating across frequency and time, we can further reduce interference. Otherwise, the performance trends are similar to those of Fig. 4. Note that idle-frame allocation does not achieve the same channel capacity as the optimal scheme because it only transmits in frames with an idle sensing result.

V. ALLOCATION FOR MULTIPLE IS USERS

The previous section derived the optimal power and transmission time allocation assuming that an orthogonal set of sub-channels had already been assigned to each IS user. This enabled us to consider each of the terminals individually and perform resource allocation based on local sensing results.

In practice, the IS base station needs to assign sub-channels to each of the IS users without knowing what the sensing outcomes will be. We therefore consider the problem of optimal sub-channel allocation based on minimizing average interference metrics. This leads to a similar formulation as compared to the average interference case in Sec. IV. Once a sub-channel allocation has been computed and fed back to the IS users, they can use the locally available sensing results to optimize their medium access. The optimal sub-channel allocation is a combinatorial problem, which is computationally more challenging than problems $P1$ through $P3$, which could be analyzed based on convex optimization. While a general analysis of this problem goes beyond the scope of this paper we find the optimal solution for fairly small problem instances by exhaustive search. A comparison with heuristic allocations suggests that efficient greedy sub-channel allocation algorithms developed for related problem setups can be adopted to this problem and yield a performance close to optimal.

A. Optimal sub-channel allocation

The problem of optimal sub-channel allocation involves assigning orthogonal sets of sub-channels to each terminal, such that the overall interference is minimized; see Fig. 6. A mathematical formulation can be based on problem $P2$. Specifically, define $f(A_i)$ as the optimal value of $P2$ where the summations over sub-channels are restricted to $n \in A_i$, i.e., the sub-channels on which a specific IS user operates. Define $f(A_i) := \infty$ if $P2$ is infeasible.

Assume that the IS base station is serving a total of $U$ users. The problem of optimal sub-channel allocation is then formulated as

$$\min_{\{A_u\}} \sum_{u=1}^U f(A_u)$$

subject to

$$A_1 \cup \cdots \cup A_U = \{1, \ldots, N\}$$

$$A_i \cap A_j = \emptyset \quad i \neq j,$$

where $A_i$ denotes the set of sub-channels assigned to terminal $i$. Due to (35) and (36) the sub-channel allocations are mutually exclusive and collectively exhaustive. Note that the above problem does not require knowledge of the sensing outcome at the individual terminals. The base station only requires knowledge of the CTMC parameters $\lambda$ and $\mu$, as well as knowledge of the IS channel coefficients $\beta$.

The above problem is difficult to solve due to its combinatorial nature and conventional sub-channel allocation methods are not easily extended to incorporate the additional dimension of allocating transmission time. For small problem instances,
however, the optimal allocation can be found by exhaustive search.

B. Suboptimal algorithm

The problem of optimal sub-channel allocation in multicarrier systems has been well-studied in the absence of transmission time allocation. Standard methods typically minimize the total transmit power subject to rate constraints. In our setup, this can be formulated mathematically as minimizing

$$\sum_{u=1}^{U} \sum_{n \in A_u} p_n$$

subject to rate and power constraints for the individual terminals. While the resulting optimization problem is still combinatorial, efficient approximation techniques have been developed with close-to-optimal performance.

Allocating sub-channels in this manner, can be used as an effective heuristic. Since good channel quality results in lower average transmission time, we conjecture that conventional sub-channel allocation may be a good approximation to the optimal interference-aware sub-channel allocation. Numerical results show that this is indeed the case in the scenarios we have examined.

C. Numerical results

Numerical results for the multi-terminal case are shown in Fig. 7. The total average transmission time overlap (summed over all IS users) is plotted with respect to the rate constraint for each individual IS users (constraints are assumed to be identical). The performance trends are the same as in the case of a single IS user. For low rate requirements we can effectively mitigate interference by assigning resources judiciously. On the other hand, as rate requirements become more stringent, there is less flexibility in accommodating the AH links. The scenario plotted in Fig. 7 corresponds to $U = 3$ terminals, $N = 5$ sub-channels, and flat Rayleigh fading IS channel coefficients.

The performance of the optimal and suboptimal sub-channel allocation schemes is very similar regardless of the rate constraint. This suggests that minimizing the total transmission power is a reasonable approximation to the optimal sub-channel allocation. In future work, we plan to corroborate this conjecture in more detail.

VI. Conclusion

In conclusion, this paper has introduced a novel cognitive coexistence framework, which enables infrastructure systems to coexist with local, ad-hoc or peer-to-peer communication links. Based on sensing and predicting the interference patterns of these ad-hoc links, the infrastructure system allocates power and transmission time judiciously such that overlaps with the ad-hoc links are minimized.

We analyzed the problem within the framework of convex optimization and derived closed-form solutions at the frame-level. These results were extended to the average rate case to reduce interference further by allocation across both frequency and time. Finally, we addressed the case of multiple IS users and provided more insight on how sub-channel allocation can be performed.

APPENDIX A

PROOF OF LEMMA 1

We denote the IS transmissions within the current frame by a finite set of closed and disjoint time intervals $I_k = [a_k, b_k]$ where each $I_k \subseteq [0, T]$ corresponds to a contiguous transmission of the IS user. Clearly, this formulation incorporates possible pauses between IS transmissions. We also require $\sum_k b_k - a_k = \rho T$, because a total of $\rho T$ transmission time needs to be allocated.

First, consider the case of an idle sensing result at the beginning of the frame, say at time $t = 0$. Then according to (3) and (1), the expected time overlap is given by

$$\frac{1}{T} \sum_k \int_{a_k}^{b_k} \Pr(X(\xi) = 1 | X(0) = 0) d\xi = \frac{1}{T} \sum_k \int_{a_k}^{b_k} \frac{\lambda}{\lambda + \mu} \left(1 - e^{-(\lambda+\mu)\xi}\right) d\xi.$$  (37)
Since the integrand is strictly increasing in $\rho$, the above expression is minimized by transmitting continuously during the time interval $[0, \rho T]$. In the case of a busy sensing result, an equivalent approach leads to a strictly decreasing integrand and therefore it is optimal to transmit during the time interval $[(1 - \rho)T, T]$ in that case.

**APPENDIX B PROOF OF LEMMA 3**

In Sec. III-B we have defined the sum power corresponding to the pair of Lagrange multipliers $[\gamma, \nu]$ as $P(\gamma, \nu)$. Further, we showed that by keeping $\gamma$ fixed and varying $\nu$ it is possible to find a $\nu^* (\gamma)$ for which the rate constraint is satisfied with equality. To simplify notation let us now define $P(\gamma)$ as the sum power associated with $[\gamma, \nu^* (\gamma)]$.

The proof that $P(\gamma)$ decreases with $\gamma$ proceeds by contradiction. First, we note that $\gamma \rightarrow 0$ implies $P(\gamma) \rightarrow \infty$ due to the structure of the optimal solutions $[\gamma_1, \nu^* (\gamma_1)]$. Assume now that $P(\gamma)$ is not monotonically decreasing. Then, because $P(\gamma)$ is continuous, there exist two different values of $\gamma$, say $\gamma_1$ and $\gamma_2$, such that $P(\gamma_1) = P(\gamma_2)$.

Based on the KKT conditions stated in Sec. III-A it is easy to verify that both $\gamma_1$ and $\gamma_2$ correspond to optimal solutions of Problem $P_1$ with rate constraint $R(\gamma_1, \nu^* (\gamma_1))$ and power constraint $P(\gamma_1, \nu^* (\gamma_1))$. Further, from the structure of the optimal solutions it is clear that the transmission time allocations associated with $\gamma_1$ and $\gamma_2$ must be different, that is, $\rho(\gamma_1) \neq \rho(\gamma_2)$. This is a contradiction, however, because Problem $P_1$ has a strictly convex objective function and therefore at most one optimal solution.

**REFERENCES**

[1] Q. Zhao and B. M. Sadler, “Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy,” IEEE Signal Process. Mag., vol. 55, no. 5, pp. 2294–2309, May 2007.

[2] L. Berlemann, C. Hoymann, G. Hiertz, and B. Walke, “Unlicensed Operation of IEEE 802.16: Coexistence with 802.11(A) in Shared Frequency Bands,” in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Sep. 2006.

[3] X. Fu, W. Ma, and Q. Zhang, “The IEEE 802.16 and 802.11a Coexistence in the License-Exempt Band,” in Proc. IEEE WCNC, Mar. 2007, pp. 1944–1949.

[4] F. H. P. Fitzek and M. D. Katz, Cognitive Wireless Networks. Springer, 2007, ch. Cellular Controlled Peer to Peer Communications: Overview and Potentials, pp. 31–59.

[5] H. Claussen, L. T. Ho, and L. G. Samuel, “An Overview of the Femtocell Concept,” Bell Labs Technical Journal, vol. 13, no. 1, pp. 221–246, 2008.

[6] J. L. Burbank and W. T. Kastch, “IEEE 802.16 Broadband Wireless Technology and Its Application to the Military Problem Space,” in Proc. IEEE Military Communications Conference (MILCOM), vol. 3, Oct. 2000, pp. 1905–1911.

[7] I. Wong and B. Evans, Resource Allocation in Multiuser Multicarrier Wireless Systems. Springer Science+Business, 2008.

[8] K. Kim, Y. Han, and S.-L. Kim, “Joint Subcarrier and Power Allocation in Uplink OFDMA Systems,” IEEE Commun. Lett., vol. 9, no. 6, pp. 526–528, Jun. 2005.

[9] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, “Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation,” IEEE J. Sel. Areas Commun., vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[10] T. Peng, W. Wang, Q. Lu, and W. Wang, “Subcarrier Allocation Based on Water-filling Level in OFDMA-based Cognitive Radio Networks,” in Proc. ACM WiCom, Sep. 2007, pp. 196–199.