

Multi-User Buffer Control with Drift Fields

Vinay Majjigi, Daniel O’Neill, Carolin Huppert and John Cioffi

Email: {vmajjigi,dconeill,cioffi}@stanford.edu, carolin.huppert@uni-ulm.de

Abstract—Multi-user buffer control is a closed-loop transmission strategy to ensure users buffers do not underflow or overflow. The dynamic scheme requires buffer state information and cooperation among transmitters. By altering the average arrival rate to users, the transmitters cause users’ buffer levels to drift away from underflow and overflow conditions. The paper presents three schemes that suggest coordination among transmitters significantly reduces the resource requirement to ensure multi-user buffer stability.

I. INTRODUCTION

Wireless systems must adapt to changing conditions caused by the dynamic nature of the environment and users’ behavior. Transmitters can adapt their power, rates, and bandwidth allocation in response to fading and application requirements, thus maximizing usage of limited resources.

Transmission strategies that only maximize spectral usage, by serving receivers in the best conditions, produce unfair allocations that poorly serve other receivers. Even a strategy that guarantees an average rate is not sufficient, as user applications may have deadline sensitive requirements. In the case where users utilize buffers, it is paramount for the transmitter to guarantee its user does not underflow and risk service interruption, e.g. in streaming video. Likewise, an overflow guarantee prevents wasting resources through dropped packets.

This work defines multiuser buffer stability as a probabilistic guarantee that all users have a minimum and a maximum number of bits in their buffer at all times. By preventing buffer underflow and overflow for users, the transmitter guarantees the short-term rate does not fluctuate wildly, and matches the user’s requirement.

Many ideas have been presented for scheduling to ensure fairness across users and short-term rate guarantees. The work of [1],[2] finds single-user capacity expressions for short-term probabilistic constraints on rate. Kelly proposed proportional fairness, a scheme that ensures a compromise between average throughput and fairness for many users [3]. The concepts of drift control for brownian systems can be found in operations research literature, and are applicable to buffer and queue management [4], [5], and [6]. Unlike much of the previous work, our approach is a closed-loop scheme that guarantees users’ buffer stability, rather than transmitter queue stability, through feedback of buffer state information (BSI) and cooperation among multiple transmitters. This variation is important given the increasing importance of streaming applications.

As an example, consider nearby femtocell access points (FAPs) each serving a user with orthogonal sets of subchannels. Consider a dynamic subset of subchannels set aside for a FAP to temporarily allocate to its user based on its buffer

state, but cycled to different FAPs as needed to ensure multi-user buffer stability. Cooperation between FAPs allows more efficient use of these subchannels by serving the weakest user.

This paper¹ considers multiple transmitters each serving a single user with an average rate and buffer stability requirement. Using a common pool of orthogonal resources, the goal is to minimize the total required resources for buffer stability. It is found that limiting user feedback by quantizing buffer state information (BSI) does not have a significant performance penalty, and cooperating transmitters require substantially less resources to ensure buffer stability than non-cooperating transmitters. Specifically, for a finite number of K users, the subset of resources set aside for buffer stability grows as $O(\sqrt{K})$ for cooperating transmitters, but grows by as much as $O(K)$ with no cooperation.

II. SYSTEM MODEL

Consider an OFDMA system with M subcarriers and K transmitters each serving a different user. The transmitters have some level of coordination such that each transmitter/user pair is allocated orthogonal subcarriers.

The subcarriers are divided into M^o subcarriers persistently allocated to users, and M^d subcarriers allocated in a dynamic manner to users, where $M = |M^o| + |M^d|$. The M^d subcarriers are viewed as the system overhead for maintaining buffer stability, see Figure 1.

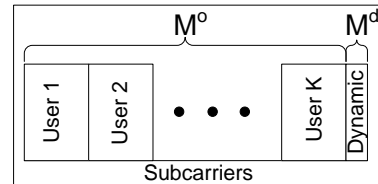


Fig. 1. K users persistently allocated M^o subcarriers. M^d dynamically allocated subcarriers.

User k has a buffer that progresses as

$$\begin{aligned} b_k[n] &= b_k[n-1] + R_k[n] - r_k[n] + \delta R_k(\vec{b}[n]) \\ &= b_k[n-1] + w_k[n] + \delta R_k(\vec{b}[n]) \end{aligned} \quad (1)$$

where $b_k[n]$ denotes the buffer occupancy at time n , and the user’s application is withdrawing a random $r_k[n]$ bits. The transmitted bits entering the buffer are split into two terms: a persistent and dynamic allocated resource delivering $R_k[n]$

¹This work was supported in part by AFOSR Complex Networks Grant FA9550-08-1-0480.

and $\delta R_k(\vec{b}[n])$ bits, respectively. And the stationary random variables are combined as $w_k[n] = R_k[n] - r_k[n]$.

The persistent resource scheme allocates M° subcarriers equally among users based on the long-term strategy $\mathbb{E}R_k = \mathbb{E}r_k$. The instantaneous rate for user k is $R_k[n]$, a random quantity based on fading and transmission policy. It then follows that $\mathbb{E}w_k = 0$. The variance of $w_k[n]$ is σ^2 .

The second rate term $\delta R_k(\vec{b}[n])$ associated with the transmitters is an event-triggered allocation of the M^d subcarriers based on users' buffer occupancy, where $\vec{b}[n] = [b_1[n] \cdots b_K[n]]^T$. In the situation with no BSI cooperation between transmitters, $\delta R_k(\vec{b}[n]) = \delta R_k(b_k[n])$.

It is assumed that $|\delta R_k(\vec{b}[n])| \ll \mathbb{E}R_k$. This condition is ensured by choosing a buffer of sufficient size so that large deviations from the nominal rate are not required. Indeed, $\delta R_k(\vec{b}[n])$ can be negative, when the transmitter is reducing the nominal rate to a user, e.g. to prevent overflow.

Buffer stability for user k is a probabilistic guarantee that its buffer does not underflow or overflow with probability ϵ_k . Underflow and overflow are both disruptive conditions for the user, for symmetry in the analysis both events are equally avoided and

$$\text{pr}(\{b_k[n] \leq 0\} \cup \{b_k[n] \geq B_{\max}\}) \leq \epsilon_k \quad (2)$$

where B_{\max} is the overflow boundary, and $B_0 = B_{\max}/2$ is the nominal operating point. Apart from the long-term average rate, (2) precludes a user from being rate starved or served excessively beyond its rate requirement, this is envisioned as a short-term rate guarantee. The system requirement ϵ^{sys} is a multi-user buffer stability guarantee that any user underflows or overflows

$$\text{pr}(\cup_{k=1}^K \{(b_k[n] \leq 0) \cup (b_k[n] \geq B_{\max})\}) \leq \epsilon^{\text{sys}} \quad (3)$$

A. Drift Fields

The buffer levels are described as a random walk process with step size $w_k[n]$. The physical interpretation, used throughout the paper, is modeling the K users' buffer levels as a particle in a K -dimensional hypercube whose size is determined by the buffer sizes. The objective for transmitters is to impart a drift field, by changing the average arrival rates to users through dynamic resource allocation, and ensure the particle does not contact the hypercube's walls.

The particle's random movement is governed by the fading channel and user application through $w_k[n]$. The drift field is modified as a transmitter varies allocated resources to a receiver and is described by $\delta \vec{R}(\vec{b}[n])$.

III. ANALYSIS

The progression of the analysis is to consider three schemes with varying amounts of user feedback and transmitter cooperation to aid the analysis and provide the result that cooperation reduces system resource requirements.

a. Continuous BSI: Transmitters have exact BSI of their user, but no BSI of other users, and apply a continuously variable drift field.

- b. Threshold BSI:* Transmitters have 1-bit of BSI, i.e. whether their user's BSI is above or below a threshold, but no BSI of other users, and apply a quantized drift field.
- c. Relative BSI:* Transmitters know relative ordering among all users, specifically, which user is closest to underflow or overflow, and apply a quantized drift field.

A. Continuous Buffer State Information

This scheme is introduced largely to provide intuition into the general dynamics of drift field schemes, and its ease in analysis. The buffer progression is modeled as a first-order auto-regressive process AR(1), where the restoring drift is proportional to the buffer level of user k ,

$$\begin{aligned} b_k[n+1] &= b_k[n] + w_k[n] - \Delta(b_k[n] - B_0) \\ &= (1 - \Delta)b_k[n] + \Delta B_0 + w_k[n] \end{aligned} \quad (4)$$

The drift field is given by $\delta R_k[n] = -\Delta(b_k[n] - B_0)$, producing a mean, or central-reverting drift towards B_0 , visualized for two users in Figure 2, where Δ is determined below. An important condition here is $0 < \Delta \ll 1$, that is the drift field is assumed to be gently nudging the process back to the middle. This assumption yields another important property, $b_k[n]$ is approximately normally distributed even though $w_k[n]$ may not be, by the Central Limit Theorem. As an AR(1) process, the steady-state variance of $b_k[n]$ is found as

$$\text{var}(b_k[n]) = \frac{\sigma^2}{2\Delta - \Delta^2} \quad (5)$$

As a mean-reverting process $\mathbb{E}b_k[n] = B_0$.

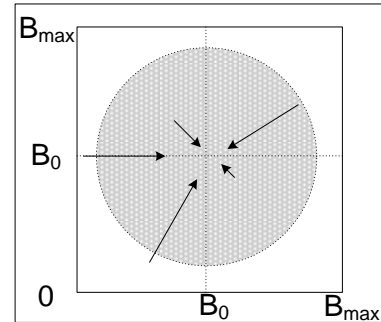


Fig. 2. Continuous BSI Scheme: Drift field is continuously variable and tends towards the center point B_0 . Shaded region is the high probability region of the particle with radius $d^{n\sigma}$.

To determine Δ for the K -user system, first consider the appropriate $\Delta^{(1)}$ for a single-user scenario, where the superscript⁽¹⁾ will stress quantities for a single user system. The buffer should not underflow or overflow with probability ϵ , and is found as the two tail probabilities of a normalized Gaussian random variable

$$\epsilon = 2Q\left(\frac{B_0 \sqrt{2\Delta^{(1)} - (\Delta^{(1)})^2}}{\sigma}\right) \quad (6)$$

The resulting $\Delta^{(1)}$ is

$$\Delta^{(1)} = 1 - \sqrt{1 - \left[\frac{\sigma Q^{-1}(\epsilon/2)}{B_0} \right]^2} \quad (7)$$

Where $Q^{-1}(\cdot)$ is the inverse Q -function of the Normal Distribution.

The step from single user to multi-user is straightforward using the Union of Events Bound to decouple users

$$\epsilon^{\text{sys}} \leq \sum_{k=1}^K \epsilon_k = K\epsilon \quad (8)$$

for $\epsilon_k = \epsilon$. This implies the K -user system requires Δ to grow as

$$\Delta = 1 - \sqrt{1 - \left[\frac{\sigma Q^{-1}(\frac{\epsilon^{\text{sys}}}{2K})}{B_0} \right]^2} \quad (9)$$

Thus, a transmitter must provide an additional average rate of $\delta R_k[n] = -\Delta(b_k[n] - B_0)$.

While the analysis for this scheme is completed, it is useful to consider the general dynamics of the particle's location to aid further analysis in Section III-B and III-C. Rather than using ϵ^{sys} to gauge stability, consider the high probability volume where the particle tends to be found.

The circular symmetry of the scheme suggests the particle tends to be in a K -hypersphere of radius $d = \sqrt{\sum_{k=1}^K (b_k - B_0)^2}$, centered at \vec{B}_0 , the K -dimensional vector with values B_0 . The time step n is dropped, and the buffer levels are given as random variables in steady-state.

As the buffer levels for each user progress independently of other users, and $b_k \sim \mathcal{N}(B_0, \frac{\sigma^2}{2\Delta - \Delta^2})$, d is given as a χ -distributed random variable with K degrees of freedom. Then

$$\mathbb{E}d = \frac{\sigma}{\sqrt{2\Delta - \Delta^2}} \frac{\sqrt{2}\Gamma((K+1)/2)}{\Gamma(K/2)} \quad (10)$$

Where $\Gamma(\cdot)$ is the Gamma function. Using Sterling's Approximation, it can be shown that $\mathbb{E}d$ grows as $O(\sqrt{K})$. Further, the standard deviation of d is approximated as,

$$\begin{aligned} \sigma_d &= \sqrt{2} \left(\frac{\sigma}{\sqrt{2\Delta - \Delta^2}} K - \frac{\sqrt{2\Delta - \Delta^2}}{\sigma} (\mathbb{E}d)^2 \right) \\ &\approx \sqrt{0.5} \frac{\sigma}{\sqrt{2\Delta - \Delta^2}} \end{aligned} \quad (11)$$

Therefore, σ_d does not depend strongly on K . Then the K -hypersphere cloud that contains the particle's likely location has radius

$$d^{n\sigma} = \mathbb{E}d + \eta\sigma_d \quad (12)$$

for $\eta > 0$. This is visualized as a circle in Figure 3 for $\eta = 3$ with plotting parameters specified in Section IV. As the radius is approximated with the mean and standard deviation, the bound is controlled by varying η and thus can be scaled to increase the probabilistic volume where the particle is found.

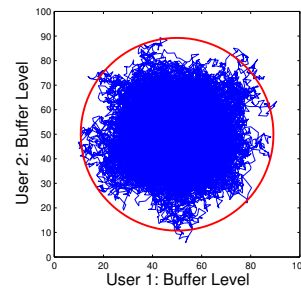


Fig. 3. Buffer sample path under Continuous Buffer State Information.

B. Threshold Buffer State Information

The Continuous BSI scheme requires an extraordinary amount of feedback from users, therefore motivates the Threshold BSI scheme that only requires 1-bit of feedback when a user's buffer level crosses a threshold, detailed in [7]. As mentioned previously, transmitters do not share their user's BSI information among other transmitters.

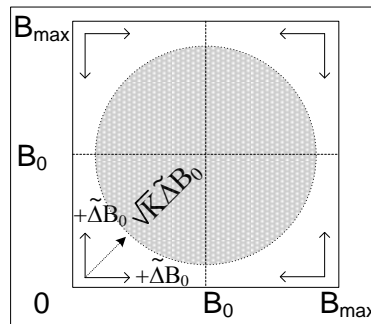


Fig. 4. Threshold BSI Scheme: Drift field is quantized based on the quadrant, and tends towards the center point \vec{B}_0 . Shaded region is the high probability region of the particle with radius $d^{n\sigma}$

In the single user case, the threshold scheme switches between two drift fields based on whether the buffer is likely to underflow or overflow. In this scheme, the restoring force takes one of two values based on whether $\tilde{b}_k[n]$, is above or below the threshold B_0

$$\tilde{b}_k[n+1] = \tilde{b}_k[n] + w_k[n] - \tilde{\Delta}B_0 \text{sign}(\tilde{b}_k[n] - B_0) \quad (13)$$

This scheme is more difficult to analyze in a steady-state manner than the continuous BSI scheme. However, as each user's buffer state is controlled independently of other users, the analysis considers a single user and then uses the Union of Events Bound to characterize its stability properties.

For a single user, the state of the system resets every time the buffer-level crosses the threshold. That is, by bounding the worst case probability of underflow/overflow after crossing a threshold to be ϵ , we can guarantee stability by at least ϵ . This guarantee was derived in [7] for a single-user underflow guarantee, and is slightly modified for an underflow/overflow guarantee as

$$\delta \tilde{R}^{(1)} = -\frac{[\frac{\sigma}{2}Q^{-1}(\epsilon/2)]^2}{B_0} \text{sign}(\tilde{b}_k[n] - B_0) \quad (14)$$

And it follows

$$\tilde{\Delta}^{(1)} = \frac{[\frac{\sigma}{2}Q^{-1}(\epsilon/2)]^2}{B_0^2} \quad (15)$$

Again, the superscript⁽¹⁾ stresses the single user system. While (15) is computed in a separate manner, it is similar to (7). Indeed, by taking the first-order Taylor series of (7) that result is exactly twice (15), a somewhat intuitive result that suggests quantizing the drift field to be half the maximum drift field required for the continuous scheme.

Using the Union of Events Bound, for K users

$$\tilde{\Delta} = \frac{[\frac{\sigma}{2}Q^{-1}(\frac{\epsilon^{\text{sys}}}{2K})]^2}{B_0^2} \quad (16)$$

Equation (16) guarantees that no user will underflow or overflow with probability ϵ^{sys} , with $\delta \tilde{R}_k[n] = -\tilde{\Delta}B_0 \text{sign}(\tilde{b}_k[n] - B_0)$, visualized for two users in Figure 4.

The analysis for this scheme is completed with (16), but the general dynamics of the scheme help motivate Section III-C. As the drift field does not always point directly to the center, the particle cloud is not circular, rather it appears to be a norm-1 ball observed through simulations as in Figure 5. To make the analysis tractable, the worst case position of the particle, which occurs at the B_0 -axes, is approximated as a random variable contained in a K -hypersphere of radius $d^{n\sigma}$ evaluated as (12).

The problem with this scheme is visualized in the lower-left quadrant in Figure 4. When all users have a low buffer state, each transmitter must allocate resources $\delta \tilde{R}_k[n] = +\tilde{\Delta}B_0$, yet affect a $\sqrt{K}\tilde{\Delta}B_0$ drift field in the K -dimensional hypercube, see Figure 4. This over-provisioning of resources for low probability events requires the transmitters to set aside $K\tilde{\Delta}B_0$ resources, and is wasteful.

The probability of all users underflowing becomes small as $K \rightarrow \infty$, indeed it can be argued that only a resource reserve of $\sqrt{K}\tilde{\Delta}B_0$ is needed in the limit. However, the concern of this paper is finite K and a required reserve of $K\tilde{\Delta}B_0$, therefore this scheme is regarded as wasteful.

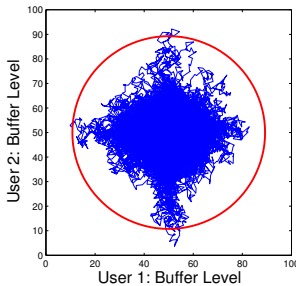


Fig. 5. Buffer sample path under Threshold Buffer State Information.

C. Relative Buffer State Information

The two schemes mentioned above are able to maintain buffer stability with a minimal amount of the drift field, thus minimize average resource usage. However, if one considers this additional drift field being created by a transmitter/receiver pair being allocated additional tones i.e. M^d subcarriers, then average resource usage is not the correct metric. Rather, from a system design perspective, minimizing the maximum number of subcarriers set aside for stability is the true optimization goal. Ideally, the dynamically allocated resources should have maximum effect, or drift, of avoiding the boundaries of the K -hypercube.

In this scheme, the buffer levels are given as

$$\hat{b}_k[n+1] = \hat{b}_k[n] + w_k[n] + \delta \hat{R}_k(\vec{b}[n]) \quad (17)$$

In a given time instant n , only user k^* is allocated additional resources. Specifically, k^* is the user closest to the boundary of the K -hypercube

$$k^* = \arg \min_k \left[(\hat{b}_k[n]), (B_{\max} - \hat{b}_k[n]) \right] \quad (18)$$

Thus user k^* is allocated resources as

$$\delta \hat{R}_{k^*}[n] = \begin{cases} +\sqrt{K}\tilde{\Delta}B_0, & \min \hat{b}_k[n] < \min(B_{\max} - \hat{b}_k[n]) \\ -\sqrt{K}\tilde{\Delta}B_0, & \min \hat{b}_k[n] \geq \min(B_{\max} - \hat{b}_k[n]) \end{cases} \quad (19)$$

That is, if the user closest to the box may underflow, the transmitter allocates resources $+\sqrt{K}\tilde{\Delta}B_0$, otherwise the user may overflow and resources are reduced by $-\sqrt{K}\tilde{\Delta}B_0$, as in Figure 6. Note $\tilde{\Delta}$ is derived from Section III-B. All other users are not allocated additional resources $\delta \hat{R}_j[n] = 0, j \neq k^*$.

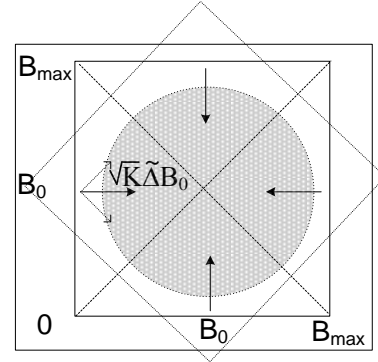


Fig. 6. Relative BSI Scheme: Drift field is quantized based on the relative BSI positions, and tends towards the center point \vec{B}_0 . Shaded region is the high probability region of the particle with radius $d^{n\sigma}$. Threshold BSI Scheme shown rotated 45° and superimposed.

The analysis of this scheme is not as straightforward as the schemes in Section III-A, III-B. The difficulty arises because the scheme does not follow the AR(1) analysis techniques of III-A, and it does not have a state-resetting mechanism as III-B. Therefore, we draw on the results and intuition from these schemes, and then validate our results in simulations.

The intuition of the relative BSI scheme is gained simply by rotating Figure 4 by 45° as shown, superimposed, in Figure

6. Here, the drift field points towards the central region, depending on the relative buffer levels of users. The transmitter allocates $\sqrt{K}\tilde{\Delta}B_0$ resources at a time, and imparts the same amount of general drift $\sqrt{K}\tilde{\Delta}B_0$. In contrast, the threshold scheme may require all users to be allocated $K\tilde{\Delta}B_0$ resources to create a drift $\sqrt{K}\tilde{\Delta}B_0$. This result suggests using the relative buffer levels of users to allocate resources.

As Figure 6 illustrates, the ϵ^{sys} stability guarantee is valid for the box tilted 45° , not the relative BSI scheme. Therefore, the stability guarantee for the relative BSI scheme is found based on the result in (12). That is, both schemes are able to contain most of the probability mass within a $d^{\eta\sigma}$ -radius hypersphere.

Through simulations, it is found that the particle is concentrated in a ∞ -norm ball, itself a hypercube cloud. The stability criteria is met and visualized in Figure 7.

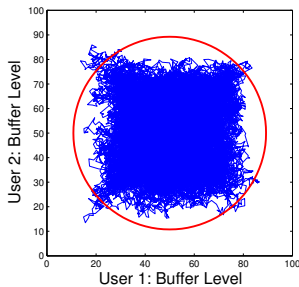


Fig. 7. Buffer sample path under Relative Buffer State Information.

IV. NUMERICAL RESULTS

In the presented simulations, $K = 15$, but for visualization, only 2 randomly selected users are plotted. Therefore Figures 3, 5, 7, and 8 are 2-Dimensional projections of the 15-Dimensional hypercube. The plotted circle has radius given by (12), for $\eta = 3$, and $\mathbb{E}d$ is computed with $K = 2$ because of the lower dimensional projection.

The simulations are 50,000 steps. The static arrival process is given by a transmitter using a water-filling strategy in Rayleigh Fading with an average SNR of 10 dB. The stability criterion, $\epsilon^{\text{sys}} = 10^{-4}$. The operating buffer level, $B_0 = 50$ bits, and the maximum level is $B_{\text{max}} = 100$ bits, and can be scaled higher for practical systems.

As is evident, with no drift field applied, i.e. an open-loop implementation, the users buffer level indiscriminately crosses the zero and B_{max} boundaries in Figure 8. The three drift field schemes are able to maintain the buffer stability criterion as designed shown in Figures 3, 5, 7.

The growth of the *peak* resources, in terms of rate, that each scheme requires is shown in Figure 9. The linear growth for required resources of the continuous and threshold schemes is compared to the square-root growth of the relative BSI scheme. Of course, not applying drift requires no additional resources.

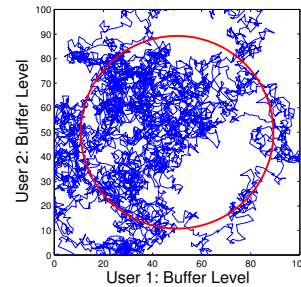


Fig. 8. Buffer sample path with no drift field applied, i.e. $\delta\vec{R}[n] = \vec{0}$.

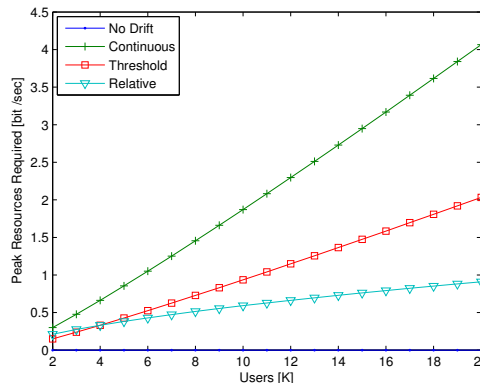


Fig. 9. Peak resources required. Linear growth with Continuous and Threshold schemes, Square-Root growth with Relative BSI scheme.

V. SUMMARY

The paper examines multi-user buffer stability with cooperating transmitters. The objective is to prevent users from underflowing or overflowing by dynamically assigning user's extra resources based on their buffer state levels. It is found that transmitters that coordinate are able to significantly reduce the peak resources required for stability.

REFERENCES

- [1] C-S. Chang and J. Thomas, "Effective bandwidth in high-speed digital networks," *IEEE Journal on Selected Areas in Comm*, vol. 13, no. 6, pp. 1091–1100, 1995.
- [2] B. Soret, M. Carmen Aguayo-Torres and J. Entrambasaguas, "Capacity with probabilistic delay constraint for Rayleigh channels," *IEEE Globecom*, 2007.
- [3] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *The Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998. [Online]. Available: <http://dx.doi.org/10.2307/3010473>
- [4] B. Ata, "Dynamic power control in a wireless static channel subject to a quality-of-service constraint," *Oper. Res.*, vol. 53, no. 5, pp. 842–851, 2005.
- [5] B. Ata, J. M. Harrison, and L. A. Shepp, "Drift rate control of a Brownian processing system," *Ann. Appl. Probab.*, vol. 15, no. math.PR/0505210. IMS-AAP-AAP-077. 2, pp. 1145–1160, May 2005.
- [6] B. Hajek, "Minimum mean hitting times of Brownian motion with constrained drift," *In Proceedings of the 27th Conference on Stochastic Processes and Their Applications*, 2001.
- [7] V. Majjigi, D. O'Neill and J. Cioffi, "Buffer state information: Two-level water-filling for fixed rate applications," *IEEE Globecom*, 2009.