Large Deviations Of Max-Weight Scheduling Policies On Convex Rate Regions

Vijay G. Subramanian
Hamilton Institute
National University of Ireland, Maynooth
Co. Kildare, Ireland.
Email: vijay.subramanian@nuim.ie

***Abstract*—We consider a single server discrete-time system with K users where the server picks operating points from a compact, convex and co-ordinate convex set in $\Re_+^K$. For this system we analyse the performance of a stablising policy that at any given time picks operating points from the allowed rate region that maximise a weighted sum of rate, where the weights depend upon the workloads of the users. Assuming a Large Deviations Principle (LDP) for the arrival processes in the Skorohod space of functions that are right-continuous with left-hand limits we establish an LDP for the workload process using a generalised version of the *contraction principle* to derive the corresponding rate function. With the LDP result available we then analyse the tail probabilities of the workloads under different buffering scenarios.**

I. Introduction

In this paper we consider a multi-class discrete-time queueing system where the server is allowed to pick operating points from within a compact convex set. Our motivation for considering compact convex rate-regions arises from information theoretic analysis of multi-user channels [1, Chapter 14] such as the multiple-access channel (MAC) or the broadcast channel (BC). Such models are particularly useful for modelling wireless systems. To operate near the capacity boundary of these channels sufficiently long code-words need to be used, which naturally leads to a discrete-time operation: pick a time long enough such that at all operating points one can choose long enough code-words such that the probability of error of decoding the code-words is small enough, and then schedule at the granularity of the chosen time-interval. For a simple class of such systems, where the rate regions are simplexes, a class of policies called the maximum weighted queue-length first policies were proposed in the context of wireless networks [2] and switches [3]. Under fairly general conditions it has been shown [2], [4], [3] that these policies are stabilising, i.e., if the average arrival rate vectors are strictly within the capacity region (in a manner to be defined later on), then the underlying Markov processes are positive recurrent [5], [6]. For a network of nodes with fixed routes for each flow, a generalisation [7] of the maximum weighted queue-length first policy where the flow with the largest weighted sum queue-length is given service over its entire route was again shown to be stabilising. A related class of policies that use the age of the head-of-the-line packet instead of queue-length have been shown to exhibit optimal performance in a Large Deviations sense over a general class of work-conserving stationary scheduling policies for a single node [8] and for a network of nodes [9] with fixed routes for each flow, in all cases the rate regions for the nodes were simplexes. In [8], [9] the queueing processes are embedded into the space of right continuous functions with left hand limits on the real line endowed with the topology of uniform convergence on compact sets. The authors then analyse the behaviour of the (stationary) maximum weighted end-to-end delay. They provide a large deviations upper-bound for the largest weighted delay first scheduling policy but only a large deviations like lower bound using inner measures over a general class of work-conserving stationary scheduling policies since the (stationary) maximum weighted delay need not be measurable for all the policies considered. For the largest weighted delay first scheduling policy the lower bound is exactly a large deviations lower bound.

The work in this paper is along the lines of the buffer overflow problem described in [10] where we consider a specific (parametric) policy and analyse its performance. Instead of just considering simplex rate-regions and two-users as in [10] we analyse a larger class of compact convex rate-regions. For these rate-regions the maximum weighted queue-length first policies can be generalised to choosing an operating point that maximises (over the rate-region) the weighted sum of rates. In keeping with the original policy we term these policies as Max-Weight policies. We prove a Large Deviations Principle (LDP) result [11] in the Skorohod space [12], [13], [14] of functions that are right-continuous with left-hand limits. Our method of proof follows the steps laid out in [15], [16], [17]. With the LDP result available we can then analyse the tail behaviour of the workloads under different buffering scenarios by an application of the *contraction principle* [11].

The paper is organised as follows. In Section II we briefly describe the theory that is needed to prove our result. Our main results, the LDP result and applications of it, are presented in Section III. The analysis proceeds by proving (in Section IV) certain properties of a deterministic problem that emerges from the limiting procedure used to prove the LDP result. We conclude in Section V with an application of the results to three two-user rate-regions: an elliptical rate-region, a Gaussian broadcast channel, a symmetrical multiple-access channel. In the interest of brevity we omit the proofs and an extensive bibliography by pointing to [18] for all the details.

II. Background Material

Here we attempt to collect together, in brief, the mathematical background material necessary to understand and prove our result. Since our paraphrasing of the material will necessarily be restrictive, for details the reader is referred to [15] (and [11, Chapter 4]), to [16], [17] for other worked examples of the method of proof, and to other references in this section. For the sake of consistency as much as possible we will use notation similar to [15], [16], [17].

Let $\mathfrak{E}$ be a metric space with metric $\rho_E(\cdot,\cdot)$. A function Π from $\mathfrak{P}(\mathfrak{E})$ the power set of $\mathfrak{E}$ to $[0,1]$ is called an idempotent probability [15, Definition 1.1.1, pp. 5-6] if $\Pi(\emptyset) = 0$, $\Pi(E) = \sup_{\mathfrak{e}\in E}\Pi(\{\mathfrak{e}\}),\ E \subseteq \mathfrak{E}$ and $\Pi(\mathfrak{E}) = 1$, and the pair $(\mathfrak{E},\Pi)$ is called an idempotent probability space. For ease of notation we will denote $\Pi(\mathfrak{e}) = \Pi(\{\mathfrak{e}\})$ for $\mathfrak{e} \in \mathfrak{E}$. A property $\mathcal{P}(\mathfrak{e}),\ \mathfrak{e} \in \mathfrak{E}$ about the elements of $\mathfrak{E}$ is defined to hold $\Pi-a.e.$ if $\Pi(\{\mathfrak{e} : \mathcal{P}(\mathfrak{e}) \text{ does not hold}\}) = 0$. A function f from a set $\mathfrak{E}$ equipped with idempotent probability Π to a set $\mathfrak{E}'$ is called an idempotent variable. The idempotent distribution of an idempotent variable f is defined as the set function $\Pi(f^{-1}(E')) = \Pi(f \in E'),\ E' \in \mathfrak{E}'$. Let $\mathcal{F}$ be a collection of subsets of $\mathfrak{E}$ that contains the null set $\emptyset$. Then Π is termed an $\mathcal{F}$-idempotent probability measure [15, Definition 1.1.1, pp. 5-6] if $\Pi(\inf_n F_n) = \inf_n \Pi(F_n)$ for every decreasing sequence of elements of $\mathcal{F}$. From now onwards unless stated otherwise we will take $\mathcal{F}_C$ to be the set of all closed sets of $\mathfrak{E}$. Define an idempotent probability measure Π to be tight if for every $\epsilon > 0$, there exists a compact $\Gamma \subseteq \mathfrak{E}$ such that $\Pi(\mathfrak{E} \setminus \Gamma) \leq \epsilon$. A tight $\mathcal{F}_C$-idempotent probability measure is called a deviability. Using [15, Lemma 1.7.4, pp. 51-52] an alternate characterisation of a deviability is a function Π such that the sets $\{\mathfrak{e} \in \mathfrak{E} : \Pi(\mathfrak{e}) \geq \gamma\}$ are compact for all $\gamma \in (0,1]$. One defines another function I from $\mathfrak{E}$ to $[0,+\infty]$ that is deemed an action functional (or good rate function) if the sets $\{\mathfrak{e} \in \mathfrak{E} : I(\mathfrak{e}) \leq x\}$ are compact for $x \in \Re_+$ and $\inf_{\mathfrak{e}\in\mathfrak{E}} I(\mathfrak{e}) = 0$ (termed lower compact). It is immediate that Π is a deviability if and only if $I(\mathfrak{e}) = -\log\Pi(\mathfrak{e})$ is an action functional. If f a mapping from $\mathfrak{E}$ to another metric space $\mathfrak{E}'$ is continuous on the sets $\{\mathfrak{e} \in \mathfrak{E} : \Pi(\mathfrak{e}) \geq \gamma\}$ for $\gamma \in (0,1]$, then $\Pi(f^{-1}(\cdot))$ is a deviability on $\mathfrak{E}'$. We define f to be a Luzin idempotent variable if $\Pi(f^{-1}(\cdot))$ is a deviability on $\mathfrak{E}'$.

Let $\{\mathbb{P}_n, n \in \mathbb{N}\}$ be a sequence of probability measures on $\mathfrak{E}$ endowed with the Borel σ-algebra, and let Π be a deviability on $\mathfrak{E}$. Let $\{m_n,\ n \in \mathbb{N}\}$ be a sequence with $m_n \to +\infty$ as $n \to +\infty$. The sequence $\{\mathbb{P}_n, n \in \mathbb{N}\}$ large deviation converges (LD converges, in short) at rate m_n to Π as $n \to +\infty$ if the inequalities $\limsup_{n\to+\infty}\mathbb{P}_n(F)^{1/m_n} \leq \Pi(F)$ and $\liminf_{n\to+\infty}\mathbb{P}_n(G)^{1/m_n} \geq \Pi(G)$ hold for all closed sets F and open sets G, respectively. Note that this is an equivalent means of describing a large deviations principle for scale m_n with good rate function $I(e) = -\log\Pi(e),\ e \in \mathfrak{E}$, given the close association of deviabilities and action functionals. Equivalent definitions that make an association with convergence of measures in the traditional sense can be found in [15, Theorem 3.1.3, pp.254-255]. As with the traditional convergence of measures the LD convergence result is shown in two steps: first, by claiming the existence of limit points by proving relative (sequential) compactness of the set of measures using some notion of tightness; and second, by demonstrating uniqueness of the limit point. A deviability Π is said to be an LD limit point of the sequence $\{\mathbb{P}_n, n \in \mathbb{N}\}$ for rate m_n if each subsequence $\{\mathbb{P}_{n_t},\ t \in \mathbb{N}\}$ contains a further subsequence $\{\mathbb{P}_{n_{t_u}},\ u \in \mathbb{N}\}$ that LD converges to Π at rate $m_{n_{t_u}}$ as $u \to +\infty$. The notion of tightness translates to the notion of exponential tightness that holds as follows: the sequence $\{\mathbb{P}_n, n \in \mathbb{N}\}$ is exponentially tight on order m_n, if for arbitrary $\epsilon > 0$ there exists a compact set $\Gamma \subseteq \mathfrak{E}$ such that $\limsup_{n\to+\infty}\mathbb{P}_n(\mathfrak{E}\setminus\Gamma)^{1/m_n} < \epsilon$. From [15, Theorem 3.1.19, pp.262-263] exponential tightness implies LD relative compactness of a sequence of measures, and therefore existence of limit points; furthermore, the limit points are all deviabilities. The LD convergence of probability measures can also be stated as the LD convergence in distribution of the associated random variables. A sequence of random variables $\{\mathfrak{X}_n,\ n \in \mathbb{N}\}$ defined on probability spaces $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$, respectively, and assuming values in $\mathfrak{E}$ LD converges at rate m_n as $n \to +\infty$ to a Luzin idempotent variable $\mathfrak{X}$ defined on idempotent probability space (Υ,Π) and assuming values in $\mathfrak{E}$ if the sequence of probability laws of $\mathfrak{X}_n$ LD converges to the idempotent distribution of $\mathfrak{X}$ at rate m_n. If the convergence of probability measures is to a deviability, then we get LD convergence in the canonical setting. The role of the *continuous mapping principle* in preserving convergence is played by the *contraction principle* [11, Theorem 4.2.1, pp. 126–127] and [15, Theorem 3.1.14 and Corollary 3.1.15 pp. 261–262] whereby if a sequence of random variables $\mathfrak{X}_n$ LD converge in distribution to $\mathfrak{X}$, and f is a ($\Pi-a.e.$) continuous function from $\mathfrak{E}$ to another metric space $\mathfrak{E}'$, then the sequence $f(\mathfrak{X}_n)$ LD converges in distribution to $f(\mathfrak{X})$ in $\mathfrak{E}'$.

In many cases of interest the LD limit points belong to a subset $\mathfrak{E}_0$ of $\mathfrak{E}$. Equip $\mathfrak{E}_0$ with the relative topology. Then we say that an idempotent probability is supported by $\mathfrak{E}_0$ if $\Pi(\mathfrak{E} \setminus \mathfrak{E}_0) = 0$. From [15, Corollary 3.1.9, pp.257-258] LD convergence of sequence $\{\mathbb{P}_n, n \in \mathbb{N}\}$ to Π that is supported by $\mathfrak{E}_0$ only needs to be checked for all $\mathfrak{E}_0$-open and $\mathfrak{E}_0$-closed Borel measurable subsets of $\mathfrak{E}$. The sequence $\{\mathbb{P}_n, n \in \mathbb{N}\}$ is termed $\mathfrak{E}_0$-exponentially tight if it is exponentially tight and every LD accumulation point Π is supported by $\mathfrak{E}_0$. Then the *contraction principle* [19] and [15, Corollary 3.1.22, pg. 264] applies to Borel-measurable but $\mathfrak{E}_0$ continuous functions.

For the results of this paper the space $\mathfrak{E}$ will be $\mathbb{D}(\mathbb{X}) := \mathbb{D}([0,1];\mathbb{X})$ the space of $\mathbb{X}$-valued, right-continuous with left-hand limits functions $\mathfrak{x} = (\mathbf{x}(t),\ t \in [0,1])$ where $\mathbb{X}$ is a complete separable metric space. Our results carry over if the setting was, instead, $\mathbb{D}([0,T];\mathbb{X})$ for fixed T with $0 < T < +\infty$. Equipping $\mathbb{D}(\mathbb{X})$ with the Skorohod J_1-topology [12], [13], [14] and metrising it with the Skorohod-Prohorov-Lindvall metric [12], [13], [14] we get a complete separable metric space; specific details can be found in [12, Chapter 3, Section 12] and [13, Sections 3.5 and 3.6].

The (closed) subset $\mathfrak{E}_0$ of $\mathbb{D}(\mathbb{X})$ that we will be dealing with

in this paper is the set of all continuous functions $\mathbb{C}(\mathbb{X}) := \mathbb{C}([0,1];\mathbb{X})$ with the induced topology, which is the uniform topology.For $\mathbf{x} \in \mathbb{X}$ define $\mathbb{C}_{\mathbf{x}}(\mathbb{X}) := \{\mathfrak{a} \in \mathbb{C}(\mathbb{X}) : \mathbf{a}(0) = \mathbf{x}\}$, a closed subset of $\mathbb{C}(\mathbb{X})$. We could prove our results for $\mathbb{D}(\mathbb{X})$ with the uniform topology. However, we prefer using the Skorohod J_1-topology since $\mathbb{D}(\mathbb{X})$ is then a complete separable metric space where by using [15, Theorem 3.1.28, pg. 268], exponential tightness is equivalent to LD relative (sequential) compactness.

Since $\mathbb{C}(\mathbb{X})$ is a closed subset of $\mathbb{D}(\mathbb{X})$ using [15, Corollaries 1.7.12 and 1.8.7] we do not distinguish between deviabilities on $\mathbb{C}(\mathbb{X})$ and deviabilities on $\mathbb{D}(\mathbb{X})$ that are supported by $\mathbb{C}(\mathbb{X})$. From [15, Remark 3.2.4 and Theorem 3.2.3, pg. 278] and [13, Sections 3.5 and 3.6] we note that a sequence of processes $\mathfrak{X}^N$ with trajectories in $\mathbb{D}(\mathbb{X})$ is $\mathbb{C}(\mathbb{X})$-exponentially tight on order N if and only if the two statements below hold.
(*Exponential tightness of random variables*) For every $t \in [0,1]$, $\mathbf{X}^N(t)$ is exponentially tight, i.e.,

$$\inf_{\Gamma \in \mathbf{\Gamma}} \limsup_{N \to +\infty} \mathbb{P}\big(\mathbf{X}^N(t) \in \mathbb{X} \setminus \Gamma\big)^{1/N} = 0, \tag{1}$$

where $\mathbf{\Gamma}$ is the set of compact subsets of $\mathbb{X}$.
(*Continuous limit points*) For every $T \in (0,1]$, $\epsilon > 0$ the following holds

$$\lim_{\delta \to 0} \limsup_{N \to +\infty} \mathbb{P}\left(\sup_{\substack{s,t \in [0,T]: \\ |s-t| \le \delta}} \rho_X\big(\mathbf{X}^N(t), \mathbf{X}^N(s)\big) > \epsilon \right)^{1/N} = 0. \tag{2}$$

If $\mathfrak{X}^N$ have trajectories in $\mathbb{D}(\Re^K)$, then by [15, Theorem 3.2.3, pg. 278] slightly simpler conditions hold to verify $\mathbb{C}(\Re^K)$-exponential tightness on order N.

For some results we will use $\mathbb{X} = \Re_+^K,\ K \in \mathbb{N}$ and for others we will use $\mathbb{X} = \Re_+^K \times \mathcal{M}(\mathcal{R}(L)),\ K,\ L \in \mathbb{N}$ where $\mathcal{M}(\mathcal{R}(L))$ is the set of all finite (non-negative) Borel measures on $\mathcal{R}(L)$ a convex compact set in $\Re_+^L$. The set of all finite (non-negative) Borel measures $\mathcal{M}(\mathfrak{E})$ on a complete separable metric space $\mathfrak{E}$, is a complete separable metric space with the Lévy-Prohorov metric [12], [13] (and [11, Theorem D.8, pp. 355-356]) and the topology of weak convergence [12], [13]. For $t \ge 0$ define $\mathcal{M}_t(\mathfrak{E}) := \{\nu \in \mathcal{M}(\mathfrak{E}) : \nu(\mathfrak{E}) \le t\}$ to be the set of (non-negative) finite measures assigning a measure at most t to $\mathfrak{E}$, and $\mathcal{M}^t(\mathfrak{E}) := \{\nu \in \mathcal{M}(\mathfrak{E}) : \nu(\mathfrak{E}) = t\}$ to be the set of (non-negative) finite measures assigning a measure exactly t to $\mathfrak{E}$. Then $\mathcal{M}_t(\mathfrak{E})$ and $\mathcal{M}^t(\mathfrak{E})$ are compact if and only if $\mathfrak{E}$ is compact [20, Section VIII.5, pg. 132] and [11, Theorem D.8, pp. 355–356]. The topology of weak convergence also results by using the Kantorovich-Wasserstein metric [21, Lemma A.1, pg. 222], [11, Theorem D.8, pp. 355-356] and [22]. Let $\mathfrak{C}^b(\mathfrak{E})$ denote the set of bounded continuous functions on $\mathfrak{E}$ that take values in $\Re$.

Finally if a sequence of random variables $\{\mathbf{X}^N,\ N \in \mathbb{N}\}$ defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assuming values in $\Re^K$ converges in probability to $\mathbf{x} \in \Re^K$ such that $\lim_{N \to +\infty} \mathbb{P}(\|\mathbf{X}_N - \mathbf{x}\| > \epsilon)^{1/N} = 0$ for all $\epsilon > 0$, then we deem the sequence as converging super-exponentially in probability at rate N and write $\mathbf{X}_N \xrightarrow{\mathbb{P}^{1/N}} \mathbf{x}$.

III. Model And A Fluid Limit

We consider a discrete-time queueing system with one server that can pick operating points from a set $\mathcal{R}(K)$ that is a compact, with non-empty interior $\mathrm{int}(\mathcal{R}(K))$, and convex subset of $\Re_+^K$ that includes the origin. We also assume that $\mathcal{R}(K)$ is coordinate-convex, i.e., if $\mathbf{r} \in \mathcal{R}(K)$, then $\hat{\mathbf{r}} \in \mathcal{R}(K)$ for all $\mathbf{0} \le \hat{\mathbf{r}} \le \mathbf{r}$ where the inequalities hold coordinate-wise [1]. Since $\mathcal{R}(K)$ is compact there exists a $r_{\max} < +\infty$ such that for every $\mathbf{r} \in \mathcal{R}(K)$ we have $r^k \le r_{\max}$ for all $k = 1, 2, \ldots, K$. For user k we assume an arrival process of work brought into the system given by a sequence $\{A_m^k\}_{m=0}^{+\infty}$ where $A_m^k \in \Re_+$ is the work brought in at time m into the queue of user k. For $-1 \le m_1 \le m_2$ integers define $A^k(m_1, m_2] := \sum_{m=m_1+1}^{m_2} A_m^k$ which is the total amount of work to arrive for user k after time slot m_1 and until time-slot m_2. Let the unfinished work in user k's queue at time $m \ge 0$ be W_m^k. Then work at time $m+1$ in the kth user's queue is given by Lindley's recursion

$$\begin{aligned} W_{m+1}^k &= \max\left(0, W_m^k - r_m^k\right) + A_m^k \\ &:= \left(W_m^k - r_m^k\right)_+ + A^k(m-1, m]. \end{aligned}$$

where $\mathbf{r}_m \in \mathcal{R}(K)$ is the operating point chosen at time m.

Our scheduling policy will be to choose a rate vector that maximises a (dynamic) weighted sum of rates over this rate region, i.e.,

$$\forall\, m \ge 1 \quad \mathbf{r}_m = \arg\max_{\mathbf{r} \in \mathcal{R}(K)} \langle \boldsymbol{\alpha}_m, \mathbf{r} \rangle$$

with components r_m^k where $\boldsymbol{\alpha}_m$ is given by

$$\alpha_m^k = \beta^k W_m^k \quad \text{s.t.} \sum_{k=1}^K \beta^k = 1,\ \beta_k > 0\ \forall k.$$

and where $< \cdot, \cdot >$ is the standard inner product in $\Re^K$. Note that $\boldsymbol{\alpha}_m$ is the Hadamard/Schur product of $\boldsymbol{\beta}$ and $\mathbf{W}_m$ and we will write this as $\boldsymbol{\alpha}_m = \boldsymbol{\beta} \circ \mathbf{W}_m$. Define the following (set-valued) functions for $\mathbf{x} \in \Re_+^K$

$$H(\mathbf{x}) := \arg\max_{\mathbf{r} \in \mathcal{R}(K)} \langle \mathbf{x}, \mathbf{r} \rangle, \tag{3}$$

$$\tilde{H}(\mathbf{x}) := H(\boldsymbol{\beta} \circ \mathbf{x}). \tag{4}$$

We will fix on a specific solution in case there is more than one maximiser. For a closed convex set $\mathcal{S} \subseteq \Re^K$ define the projection of element $x \in \Re^K$ to be the unique element $\mathbf{x}^* \in \mathcal{S}$ that solves

$$\min_{\mathbf{y} \in \mathcal{S}} \|\mathbf{x} - \mathbf{y}\|^2,$$

where $\|\cdot\|$ is the Euclidean norm given by $\sqrt{< \mathbf{x}, \mathbf{x} >}$ for $\mathbf{x} \in \Re_+^K$. We define the function from $\mathbf{x}$ to $\mathbf{x}^*$ for a given $\mathcal{S}$ to be $\mathrm{Proj}_{\mathcal{S}}(\mathbf{x})$. For every $\mathbf{x} \in \Re^K$ is clear that $\tilde{H}(\mathbf{x})$ is a closed and

[1] Henceforth, unless specified otherwise, we assume that all vector inequalities hold coordinate-wise.

convex set. Then the specific operating point that we choose at time m is given by

$$\mathbf{r}_m = \mathrm{Proj}_{\tilde{H}(\mathbf{W}_m)}(\mathbf{0}).$$

Based upon the above definition we call operating point $\mathbf{r}_m$ the minimum norm solution.

Using the operating point $\mathbf{r}_m$ at time m we get

$$\begin{aligned} W^k_{m+1} &= \left(W^k_m - r^k_m\right)_+ + A^k(m-1,m] \\ &= W^k_m - \min(W^k_m, r^k_m) + A^k(m-1,m] \\ &= W^k_0 - \sum_{l=0}^{m} S^k_l + A^k(-1,m] \\ &= W^k_0 - S^k(-1,m] + A^k(-1,m], \end{aligned} \tag{5}$$

where for $m \geq 1$ we define $S^k_m := \min(W^k_m, r^k_m) \leq r_{\max}$, which is the amount of work from the queue of user k served at time m. Coordinate convexity ensures that for every $\mathbf{r} \in \mathcal{R}(K)$ every point $\min(\mathbf{x}, \mathbf{r})$ belongs to $\mathcal{R}(K)$ as $\mathbf{x}$ is allowed to vary in $\Re^K_+$.

Assume that we are given a sequence $\{\mathbf{W}^N_0\}_{N\in\mathbb{N}}$ taking values in $\Re^K_+$ that accounts for the vector of initial work in the system. We then embed the sequences $\{A^k(-1,m]\}$, $\{S^k(-1,m]\}$ and $\{W^k_m\}$ into functions in $\mathbb{D}(\Re^K)$ by defining (scaling both space and time) for $t \in [0,1]$ the following: $\tilde{A}^{k,N}(t) := \frac{A^k(-1,\lfloor Nt\rfloor]}{N}$, $\tilde{S}^{k,N}(t) := \frac{S^{k,N}(-1,\lfloor Nt\rfloor]}{N}$ and $\tilde{W}^{k,N}(t) := \frac{W^{k,N}_{\lfloor Nt\rfloor}}{N}$ for $N \in \mathbb{N}$ where $\lfloor t\rfloor$ is the largest integer less than or equal to t. The index N in $S^{k,N}$ and $W^{k,N}$ takes into account the different initial workload vector given by $\mathbf{W}^N_0$. Denote the vector quantities by $\tilde{\mathbf{A}}^N(t)$, $\tilde{\mathbf{S}}^N(t)$ and $\tilde{\mathbf{W}}^N(t)$, respectively. Also define the processes $\mathfrak{A}^N := (\tilde{\mathbf{A}}^N(t),\ t \in [0,1])$, $\mathfrak{S}^N := (\tilde{\mathbf{S}}^N(t),\ t \in [0,1])$, and $\mathfrak{W}^N := (\tilde{\mathbf{W}}^N(t),\ t \in [0,1])$. The workload arrivals sequence $\{\mathbf{A}_m\}_{m\in\mathbb{N}}$ and the initial workload vector sequence $\{\mathbf{W}^N_0\}_{N\in\mathbb{N}}$ are assumed to be defined on a common complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Construct the random empirical measure $\tilde{\Psi}(-1,m](\cdot) := \sum_{l=0}^{m} \boldsymbol{\delta}_{\mathbf{r}_m}(\cdot)$ where $\boldsymbol{\delta}_{\mathbf{x}}(\cdot)$ is the Dirac measure. Define the scaled empirical measure process $\Psi^N(t) := \frac{\tilde{\Psi}^N(-1,\lfloor Nt\rfloor]}{N}$ for $t \in [0,1]$. Again the index N accounts for the different initial workload vector. Let $\mathcal{M}(\mathcal{R}(K))$ be the set of finite (non-negative) Borel measures on $\mathcal{R}(K)$; when endowed with the topology of weak convergence of measures generated by the Kantorovich-Wasserstein metric, $\mathcal{M}(\mathcal{R}(K))$ is a complete separable metric space. Then the processes $\boldsymbol{\Psi}^N$ take values in $\mathbb{D}(\mathcal{M}(\mathcal{R}(K)))$ again with the Skorohod J_1 topology [13]. In fact for every $t \in [0,1]$ we have $\Psi^N(t) \in \mathcal{M}_{t+1}(\mathcal{R}(K)) = \{\nu \in \mathcal{M}(\mathcal{R}(K)) :\ \nu(\mathcal{R}(K)) \leq t+1\}$ where $\mathcal{M}_{t+1}(\mathcal{R}(K))$ is a compact subset of $\mathcal{M}(\mathcal{R}(K))$. For a Borel measurable function f from $\mathcal{R}(K)$ to $\Re$ denote the integral (if it exists) with respect to a measure $\nu \in \mathcal{M}(\mathcal{R}(K))$ by $\int_{\mathcal{R}(K)} f d\nu$; this is a random variable taking values in $\Re$ that we denote as $< \nu, f >$.

For our convergence proofs we will be considering processes $(\mathfrak{A}^N, \mathfrak{S}^N, \boldsymbol{\Psi}^N, \mathfrak{W}^N)$ taking values in the Skorohod space $\mathbb{D}(\Re^K \times \Re^K \times \mathcal{M}(\mathcal{R}(K)) \times \Re^K)$; we denote the complete separable metric space $\Re^K \times \Re^K \times \mathcal{M}(\mathcal{R}(K)) \times \Re^K$ by $\mathbb{X}$. Denote the Euclidean metric on $\Re^K$ by ρ_E and the Kantorovich-Wasserstein metric on $\mathcal{M}(\mathcal{R}(K))$ by ρ_{KL}. Then for two elements $(\mathbf{a}_1, \mathbf{s}_1, \Phi_1, \mathbf{w}_1)$, $(\mathbf{a}_2, \mathbf{s}_2, \Phi_2, \mathbf{w}_2) \in \mathbb{X}$ the distance between the two elements is given by the following metric

$$\begin{aligned} &\rho_X\big((\mathbf{a}_1, \mathbf{s}_1, \Phi_1, \mathbf{w}_1), (\mathbf{a}_2, \mathbf{s}_2, \Phi_2, \mathbf{w}_2)\big) := \\ &\quad \max(\rho_E(\mathbf{a}_1, \mathbf{a}_2), \rho_E(\mathbf{s}_1, \mathbf{s}_2), \rho_{KL}(\Phi_1, \Phi_2), \rho_E(\mathbf{w}_1, \mathbf{w}_2)). \end{aligned}$$

The topology that results from this metric is the product topology.

To prove the required large deviations result we will follow the programme outlined in [15], [16], [17]. Loosely speaking, we first show in Theorem 3.1 that the sequence of measures on a metric space $\mathfrak{E}$ is large deviations relatively compact using exponential tightness. In proving relatively compactness we will prove that all the limit points are supported on $\mathfrak{E}_0$ a closed subset of $\mathfrak{E}$. The limit points are determined by weak solutions to idempotent equations the properties of which are characterised by taking large deviation limits of the stochastic equations that determine the behaviour of the original system. Compactness of the rate-regions plays an important role in proving exponential tightness and characterising the large deviations limit points. Using all the characterised properties of the idempotent equations we show the existence of unique solutions to the idempotent equations in Theorem 4.1 leading to an LDP result in Theorem 3.2. The convexity of the rate-regions and the nature of the scheduling policy (maximising a linear functional over a convex set) play an important part in not only proving the existence and uniqueness of solutions to the idempotent equations but also in providing a simple expression for the solution. The LDP result then follows directly from Theorem 4.1 by appealing to a generalised version of the *contraction principle* [19] and [15, Corollary 3.1.22, pg. 264].

Assume we are given a function $\chi^{\mathbf{A}} : \Re^K_+ \to [0, +\infty]$ that attains zero at some $\boldsymbol{\mu} \in \Re^K_+$ such that $\boldsymbol{\mu} < \boldsymbol{\lambda}^*$ with $\boldsymbol{\lambda}^* \in \mathcal{R}(K)$ where the inequality holds coordinate-wise. This is used to define function $\mathbf{I}^{\mathfrak{A}} : \mathbb{D}(\Re^K_+) \to \Re_+$ by

$$\mathbf{I}^{\mathfrak{A}}(\mathfrak{a}) = \int_0^1 \chi^{\mathbf{A}}\left(\dot{\mathbf{a}}(t)\right) dt$$

if the function $\mathfrak{a} = \big(\mathbf{a}(t),\ t \in [0,1]\big) \in \mathbb{D}(\Re^K_+)$ is absolutely continuous such that $\mathbf{a}(0) = \mathbf{0}$, and where $\dot{\mathfrak{a}} \in \mathcal{L}^1\big([0,1]; \Re^K\big)$ is the (Lebesgue) *a.e.* derivative of $\mathfrak{a}$ taking values in $\Re^K_+$; $\mathbf{I}^A(\mathfrak{a})$ is defined to be equal to $+\infty$ if the function $\mathfrak{a} \in \mathbb{D}(\Re^K_+)$ does not have the above properties. For $\mathbf{x} \in \Re^K_+$ we define by $\mathbb{AC}_{\mathbf{x}}$ the set of absolutely continuous functions in $\mathbb{D}(\Re^K_+)$ such that for $\mathfrak{a} \in \mathbb{AC}_{\mathbf{x}}$ we have $\mathbf{a}(0) = \mathbf{x}$, and where $\dot{\mathfrak{a}} \in \mathcal{L}^1([0,1]; \Re^K)$ takes values in $\Re^K_+$. It is assumed that $\mathbf{I}^{\mathfrak{A}}(\cdot)$ is an action functional on $\mathbb{D}(\Re^K_+)$ which in turn implies that $\chi^{\mathbf{A}}(\cdot)$ is an action functional on $\Re^K_+$. The function $\chi^{\mathbf{A}}(\cdot)$ being convex and lower compact is sufficient [21, Lemma 8, pg. 203] for $\mathbf{I}^{\mathfrak{A}}(\cdot)$ to be an action functional.

We assume that the arrival process of workloads is such that $\{\mathfrak{A}^N,\ N \in \mathbb{N}\}$ satisfies an LDP at rate N as $N \to +\infty$ in the space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{A}}(\cdot)$. If the arrival process of workloads is such that $\{\mathbf{A}_m\}_{m\geq 0}$ is a sequence of *i.i.d.* $\Re_+^K$ valued random variables with $\mathbb{E}\left(e^{<\mathbf{x},\mathbf{A}_0>}\right) < +\infty$ for all $\mathbf{x} \in \Re^K$ where $< \cdot, \cdot >$ is the inner product on $\Re^K$, then by Mogulskii's theorem [23, Theorem 2.15, pg. 25] and [24], [25] we have $\{\mathfrak{A}^N,\ N \in \mathbb{N}\}$ satisfying an LDP at rate N as $N \to +\infty$ in the space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{A}}(\cdot)$ with

$$\chi^{\mathbf{A}}(\mathbf{x}) = \sup_{\mathbf{y}\in\Re^K}\left(\left\langle \mathbf{y},\mathbf{x}\right\rangle - \log \mathbb{E}\left(e^{\left\langle \mathbf{y},\mathbf{A}_0\right\rangle}\right)\right) \quad \forall\, \mathbf{x} \in \Re_+^K. \tag{6}$$

See [21, Theorem 5, pg. 216] for general mixing conditions on a stationary sequence $\{\mathbf{A}_m\}$ so that $\{\mathfrak{A}^N,\ N \in \mathbb{N}\}$ satisfies an LDP at rate N as $N \to +\infty$ in the space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{A}}(\cdot)$.

Under the above assumptions about the arrival processes we now prove the $\mathbb{C}(\mathbb{X})$-exponential tightness of the sequence $(\mathfrak{A}^N, \mathfrak{S}^N, \mathbf{\Psi}^N, \mathfrak{W}^N)$.

Theorem 3.1: Assume that the arrival process is such that the sequence $\{\mathfrak{A}^N,\ N \in \mathbb{N}\}$ satisfies an LDP at rate N as $N \to +\infty$ in the space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{A}}(\cdot)$. Also assume[2] that $\frac{\mathbf{W}_0^N}{N} \xrightarrow{\mathbb{P}^{1/N}} \mathbf{w}(0)$. Then the sequence $(\mathfrak{A}^N, \mathfrak{S}^N, \mathbf{\Psi}^N, \mathfrak{W}^N)$ is $\mathbb{C}(\mathbb{X})$-exponentially tight on order N in $\mathbb{D}(\mathbb{X})$.

If an idempotent process $(\mathfrak{a}, \mathfrak{s}, \mathbf{\Phi}, \mathfrak{w})$ defined on an idempotent probability space (Υ, Π) and having trajectories in $\mathbb{C}(\mathbb{X})$ is a limit point of $(\mathfrak{A}^N, \mathfrak{S}^N, \mathbf{\Psi}^N, \mathfrak{W}^N)$ for LD convergence in distribution at rate N, then the following properties hold $\Pi-a.e.$ for any limit point $(\mathfrak{a}, \mathfrak{s}, \mathbf{\Phi}, \mathfrak{w})$:

(i) the function $\mathfrak{a}$ is component-wise non-negative and non-decreasing, and absolutely continuous with $\mathbf{a}(0) = \mathbf{0}$;

(ii) the function $\mathfrak{s}$ is component-wise non-negative and non-decreasing, and Lipschitz continuous with $\mathbf{s}(0) = \mathbf{0}$.

(iii) the measure-valued function $\mathbf{\Phi}$ is such that $\Phi(t) \in \mathcal{M}^t(\mathcal{R}(K))$ for all $t \in [0,1]$, absolutely continuous with respect to the total variation norm [26, pg. 35-38, 118-119] such that $\Phi(t) - \Phi(u) \in \mathcal{M}^{t-u}(\mathcal{R}(K))$ for all $1 \geq t > u \geq 0$ with $\Phi(0)(\mathcal{R}(K)) = 0$, and possesses a weak derivative[3] $\dot{\Phi}(t)$ for almost every $t \in [0,1]$. The following inequality holds for all $t, u \in [0,1]$ with $t \geq u$

$$s^k(t) - s^k(u) \leq \left\langle \Phi(t) - \Phi(u), e_k\right\rangle, \tag{7}$$

where $e_k : \mathcal{R}(K) \to \Re_+$ is the kth-coordinate projection operator such that $e_k(\mathbf{r}) = r_k$;

(iv) the function $\mathfrak{w}$ is component-wise non-negative, and absolutely continuous with $\mathbf{w}(0)$ given such that

$$\dot{\mathbf{w}}(t) = \dot{\mathbf{a}}(t) - \dot{\mathbf{s}}(t) \quad \text{for (Lebesgue) a.e. } t \in [0,1]; \tag{8}$$

(v) if $w^k(t) > 0$ for $t \in [t_1, t_2]$ for $t_1,\ t_2 \in [0,1]$, then it follows that

$$s^k(t) - s^k(u) = \left\langle \Phi(t) - \Phi(s), e_k\right\rangle \tag{9}$$

for all $t \geq u$ with $t,\ u\ \in [t_1, t_2]$; and

(vi) for (Lebesgue) almost every $t \in [0,1]$,

$$\dot{\Phi}(t)\left(\mathcal{R}(K) \setminus \tilde{H}(\mathbf{w}(t))\right) = 0. \tag{10}$$

Thus $\Pi-a.e.$ every limit point is an absolutely continuous solution of the following differential inclusion for all $t \in [0,1]$:

$$\dot{\mathbf{w}}(t) \in \dot{\mathbf{a}}(t) - \tilde{H}(\mathbf{w}(t)) \tag{11}$$

with $\mathbf{w}(0)$ the initial condition.

Proof: Refer to [18]. ■

We will show the existence of solutions of (11), and show uniqueness of the solution and other properties as a consequence of Theorem 4.1 in Section IV.

Remark: If we make the additional assumption that $\chi^{\mathbf{A}}(\boldsymbol{\nu}) > 0$ for all $\boldsymbol{\nu} \in \Re_+^K \setminus \{\boldsymbol{\mu}\}$ and $\{\mathbf{A}_m\}_{m\geq 0}$ being stationary, then we can construct a (regular) fluid limit (a functional strong law of large numbers result) that will obey an relationship similar to (11) given by absolutely continuous solutions to

$$\dot{\mathbf{w}}(t) \in \boldsymbol{\mu} - \tilde{H}(\mathbf{w}(t)) \tag{12}$$

with $\mathbf{w}(0)$ the initial condition such that $\sum_{k=1}^K w^k(0) = 1$. Now one can easily argue for stability [4], [5], [6] (existence of stationary regime and stationary distribution) by using a quadratic Lyapunov function $V(t) := \frac{\|\sqrt{\boldsymbol{\beta}} \circ \mathbf{w}(t)\|^2}{2}$ under the conditions of [21, Theorem 5, pg. 216]. The details are very similar to those in [4] and are skipped for brevity.

Now we state our main result

Theorem 3.2: Assume that the sequence of arrival processes $\{\mathfrak{A}^N,\ N \in \mathbb{N}\}$ satisfy an LDP at rate N as $N \to +\infty$ in the space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{A}}(\cdot)$. Also assume that $\frac{\mathbf{W}^N(0)}{N} \xrightarrow{\mathbb{P}^{1/N}} \mathbf{w}(0)$. Then the sequence $\mathfrak{W}^N$ obeys an LDP for scale N in the Skorohod space $\mathbb{D}(\Re_+^K)$ with action functional $\mathbf{I}^{\mathfrak{W}}_{\mathbf{w}(0)}(\cdot)$.

We defer the proof and the identification of the action functional to Section IV. We can immediately write down a corollary to Theorem 3.2 that considers many applications of the result.

Corollary 1: Under the conditions of Theorem 3.2 with $\mathbf{w}(0) = \mathbf{0}$, for $\mathbf{x} \in \Re_+^K$, $x \in \Re_+$ and $t \in [0,1]$ we have

$$\limsup_{N\to+\infty} \frac{\log\left(\mathbb{P}(\mathbf{W}(\lfloor Nt\rfloor) \geq N\mathbf{x})\right)}{N} \leq -\inf_{\mathbf{y}\in\Re_+^K : \mathbf{y}\geq\mathbf{x}} \mathbf{J}(\mathbf{y}, t), \tag{13}$$

$$\liminf_{N\to+\infty} \frac{\log\left(\mathbb{P}(\mathbf{W}(\lfloor Nt\rfloor) > N\mathbf{x})\right)}{N} \geq -\inf_{\mathbf{y}\in\Re_+^K : \mathbf{y}>\mathbf{x}} \mathbf{J}(\mathbf{y}, t),, \tag{14}$$

[2]This can be relaxed with a suitable LDP assumption - LDP with good rate function, for example.

[3]Following [21] by weak derivative we mean the (weak) convergence of $\frac{\Phi(t+\epsilon)-\Phi(t)}{\epsilon} \in \mathcal{M}(\mathcal{R}(K))$ as $\epsilon \to 0$ with the limit denoted as $\dot{\Phi}(t)$. For absolutely continuous $\Phi(t)$ we have $\Phi(t) - \Phi(u) = \int_u^t \dot{\Phi}(\tau)d\tau$ where the integral is interpreted set-wise, i.e., for $C \in \mathcal{B}(\mathcal{R}(K))$ we have $\big(\Phi(t) - \Phi(u)\big)\Big(C\Big) = \int_u^t \dot{\Phi}(\tau)(C)d\tau$.

where

$$\mathbf{J}(\mathbf{x},t) := \inf_{\mathfrak{w}\in\mathbb{C}_{\mathbf{0}}(\Re_+^K):\mathbf{w}(t)=\mathbf{x}} \mathbf{I}_{\mathbf{0}}^{\mathfrak{W}}(\mathfrak{w}). \tag{15}$$

Furthermore, we also have

$$\limsup_{N\to+\infty} \frac{\log\big(\mathbb{P}(\max_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) \geq Nx)\big)}{N} \leq -\min_{k=1,2,\ldots,K} \inf_{\mathbf{y}\in\Re_+^K:y^k\geq x} \mathbf{J}(\mathbf{y},t), \tag{16}$$

$$\liminf_{N\to+\infty} \frac{\log\big(\mathbb{P}(\max_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) > Nx)\big)}{N} \geq -\min_{k=1,2,\ldots,K} \inf_{\mathbf{y}\in\Re_+^K:y^k> x} \mathbf{J}(\mathbf{y},t). \tag{17}$$

$$\limsup_{N\to+\infty} \frac{\log\big(\mathbb{P}(\sum_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) \geq Nx)\big)}{N} \leq -\inf_{\mathbf{y}\in\Re_+^K:\sum_{k=1}^K y^k\geq x} \mathbf{J}(\mathbf{y},t), \tag{18}$$

$$\liminf_{N\to+\infty} \frac{\log\big(\mathbb{P}(\sum_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) > Nx)\big)}{N} \geq -\inf_{\mathbf{y}\in\Re_+^K:\sum_{k=1}^K y^k> x} \mathbf{J}(\mathbf{y},t). \tag{19}$$

$$\limsup_{N\to+\infty} \frac{\log\big(\mathbb{P}(\max_{k=1,2,\ldots,K} \sup_{t\in[0,1]} W^k(\lfloor Nt\rfloor) \geq Nx)\big)}{N} \leq -\min_{k=1,2,\ldots,K} \inf_{\mathbf{y}\in\Re_+^K:y^k\geq x} \inf_{t\in[0,1]} \mathbf{J}(\mathbf{y},t), \tag{20}$$

$$\liminf_{N\to+\infty} \frac{\log\big(\mathbb{P}(\max_{k=1,2,\ldots,K} \sup_{t\in[0,1]} W^k(\lfloor Nt\rfloor) > Nx)\big)}{N} \geq -\min_{k=1,2,\ldots,K} \inf_{\mathbf{y}\in\Re_+^K:y^k> x} \inf_{t\in[0,1]} \mathbf{J}(\mathbf{y},t). \tag{21}$$

$$\limsup_{N\to+\infty} \frac{\log\big(\mathbb{P}(\sup_{t\in[0,1]} \sum_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) \geq Nx)\big)}{N} \leq -\inf_{\mathbf{y}\in\Re_+^K:\sum_{k=1}^K y^k\geq x} \inf_{t\in[0,1]} \mathbf{J}(\mathbf{y},t), \tag{22}$$

$$\liminf_{N\to+\infty} \frac{\log\big(\mathbb{P}(\sup_{t\in[0,1]} \sum_{k=1,2,\ldots,K} W^k(\lfloor Nt\rfloor) > Nx)\big)}{N} \geq -\inf_{\mathbf{y}\in\Re_+^K:\sum_{k=1}^K y^k> x} \inf_{t\in[0,1]} \mathbf{J}(\mathbf{y},t). \tag{23}$$

Proof: Refer to [18]. ■

The restriction $\mathbf{w}(0) = \mathbf{0}$ is only to simplify our further characterisation of $\mathbf{J}(\mathbf{x},t)$. Note that (20) and (21) are useful in calculating the tail probabilities of the workload when each user has a buffer to itself and (22) and (23) are useful when there is a shared buffer.

IV. Analysis of Fluid Limit

Before addressing the main result of this section we state and prove a few preliminary results that will be key for the analysis of the fluid limit. We are interested in the properties of $H(\mathbf{x}) = \arg\max_{\mathbf{r}\in\mathcal{R}(K)} \langle \mathbf{x},\mathbf{r}\rangle$ and $\tilde{H}(\mathbf{x}) = \arg\max_{\mathbf{r}\in\mathcal{R}(K)} \langle \boldsymbol{\beta}\circ\mathbf{x},\mathbf{r}\rangle$ for $\mathbf{x}\in\Re^K$.

Let $\mathfrak{X}$ be a Hilbert space. A set-valued map $\mathcal{H}$ from $\mathfrak{X}$ to $\mathfrak{P}(\mathfrak{X})$ (the power set of $\mathfrak{X}$) with domain $\mathrm{Dom}(\mathcal{H})$ is *monotone* ([27, Definition 2.1, pg. 20] and [28, Chapter 12]) if and only if

$$\forall \mathfrak{x}_1,\ \mathfrak{x}_2 \in \mathrm{Dom}(F),\ \forall \mathfrak{v}_i \in \mathcal{H}(\mathfrak{x}_i),\ i=1,2,$$
$$\langle \mathfrak{v}_1-\mathfrak{v}_2, \mathfrak{x}_1-\mathfrak{x}_2\rangle \geq 0, \tag{24}$$

where $< \cdot,\cdot >$ is the inner-product on $\mathfrak{X}$. A *monotone* set-valued map $\mathcal{H}$ is *maximal* ([27, Definition 2.2, pg. 22] and [28, Chapter 12]) if there is no other monotone set-valued map $\tilde{\mathcal{H}}$ whose graph strictly contains the graph of $\mathcal{H}$. The reader is referred to [27], [29], [30], [28] for the properties of monotone maps, maximal monotone maps and their connections to convex analysis, functional analysis and semigroups of non-expansive maps.

Lemma 1: $H(\mathbf{x}) = \arg\max_{\mathbf{r}\in\mathcal{R}(K)} \langle \mathbf{x},\mathbf{r}\rangle$ and $\tilde{H}(\mathbf{x}) = \arg\max_{\mathbf{r}\in\mathcal{R}(K)} \langle \boldsymbol{\beta}\circ\mathbf{x},\mathbf{r}\rangle$ for $\mathbf{x}\in\Re^K$ are maximal monotone maps from $\Re^K$ to $\mathcal{R}(K)\subset\Re^K$.

Proof: Refer to [18]. ■

As described in Section III each policy in the class of Max-Weight scheduling policies can be associated with a unique vector $\tilde{\boldsymbol{\beta}}\in\Re_+^K$ such that $\tilde{\boldsymbol{\beta}} > \mathbf{0}$ and $\sum_{k=1}^K \tilde{\beta}^k = 1$. We can demonstrate [18] that the performance of a Max-Weight policy with a given $\tilde{\boldsymbol{\beta}}$ can be quantified by analysing the performance of a Max-Weight policy with weights $\boldsymbol{\beta} = \underbrace{\left(\frac{1}{K},\frac{1}{K},\ldots,\frac{1}{K}\right)}_{K \text{ times}}$ but with a new rate-region that is a scaled version of the original rate region. Thus, without loss of generality, from now onwards we assume that $\boldsymbol{\beta} = \underbrace{\left(\frac{1}{K},\frac{1}{K},\ldots,\frac{1}{K}\right)}_{K \text{ times}}$ and analyse the solutions of the following differential inclusion

$$\forall t\in[0,1] \quad \dot{\mathbf{w}}(t) \in \dot{\mathbf{a}}(t) - H(\mathbf{w}(t)), \tag{25}$$

where $\mathfrak{a}\in\mathbb{AC}_{\mathbf{0}}$, $\mathbf{w}(0)\in\Re^K$ is given, and $H(\cdot):\Re^K\to\mathfrak{P}\left(\Re^K\right)$ is a *maximal monotone* set-valued map.

Instead of just seeking a solution to (25) when $\mathfrak{a}\in\mathbb{AC}_{\mathbf{0}}$, in [31] (weak) solutions for $\mathfrak{a}\in\mathbb{C}_{\mathbf{x}(0)}(\Re_+^K)$ were analysed. Using the results of [31] we then have the following theorem where we define $\mathrm{cl}(\mathcal{S})$ to be the closure of set $\mathcal{S}$ and the (*a.e.*) right derivative at time $t\in[0,1]$ of an absolutely continuous function $\mathfrak{w}\in\mathbb{AC}$ to be $\frac{d^+\mathfrak{w}}{dt}(t)$.

Theorem 4.1: Let $\mathcal{H}$ be a maximal monotone map such that its domain $\mathrm{Dom}(\mathcal{H})$ has a non-empty interior $\mathrm{int}(\mathrm{Dom}(\mathcal{H}))$. If $\mathbf{w}(0)\in\mathrm{cl}(\mathrm{Dom}(\mathcal{H}))$, then for $\mathfrak{a}\in\mathbb{C}_{\mathbf{0}}(\Re_+^K)$ there exists a unique (weak) solution of

$$\dot{\mathbf{w}}(t) \in \dot{\mathbf{a}}(t) - \mathcal{H}(\mathbf{w}(t)) \quad \forall t\in[0,1], \tag{26}$$

with $\mathfrak{w}\in\mathbb{C}_{\mathbf{w}(0)}(\Re_+^K)$ taking values in $\mathrm{cl}(\mathrm{Dom}(\mathcal{H}))$. The map from $\mathbf{w}(0)+\mathfrak{a}$ to $\mathfrak{w}$ is continuous with the uniform topology on $\mathbb{C}(\Re_+^K)$. If, in addition, $\mathfrak{a}\in\mathbb{AC}_{\mathbf{0}}$, then $\mathfrak{w}\in\mathbb{AC}_{\mathbf{w}(0)}$ (and is a strong solution) with right derivative $\frac{d^+\mathfrak{w}}{dt}$ given by

$$\frac{d^+\mathbf{w}}{dt}(t) = \dot{\mathbf{a}}(t) - \mathrm{Proj}_{\mathcal{H}(\mathbf{w}(t))}\left(\dot{\mathbf{a}}(t)\right). \tag{27}$$

Proof: Refer to [18]. ■

Now we spell out the details of the proof of Theorem 3.2.

Proof of Theorem 3.2: First note that $H(\cdot)$ is *maximal monotone* with $\mathrm{Dom}(H) = \Re^K$. Now if one considers the function in $\mathbb{C}_{\mathbf{x}(0)}(\Re_+^K)$ given by $\hat{\mathfrak{a}} = \mathbf{w}(0) + \mathfrak{a}$ for $\mathfrak{a} \in \mathbb{C}_{\mathbf{0}}(\Re_+^K)$, then using the results of Cépa [31] summarised in Theorem 4.1 one can show the existence of a unique continuous (weak) solution to (25) for continuous input $\hat{\mathfrak{a}}$ such that the map from $\hat{\mathfrak{a}}$ to $\mathfrak{w}$ is continuous with the uniform topology on $\mathbb{C}(\Re_+^K)$. Thus we can directly apply the general version of the contraction principle [19] and [15, Corollary 3.1.22, pg. 264] to prove the LD convergence result since the composite map from $\mathfrak{a} \in \mathbb{C}_{\mathbf{0}}(\Re_+^K)$ to $\mathfrak{w}$ through $\mathbf{w}(0) + \mathfrak{a}$ is, by consequence, continuous with the uniform topology of $\mathbb{C}(\Re_+^K)$. Define the composite map to be $\mathcal{T}$.

The action functional is also immediate now. Let $\mathcal{T}_A$ be the image under $\mathcal{T}$ of absolutely continuous functions $\mathfrak{a} \in \mathbb{AC}_{\mathbf{0}}$. Note that $\mathcal{T}_A$ is a subset of the set of absolutely continuous functions from $[0,1]$ with initial value $\mathbf{w}(0)$. Then the action functional for the LDP result $\mathbf{I}^{\mathfrak{W}}_{\mathbf{w}(0)}(\cdot)$ is given as follows: if $\mathfrak{w} \in \mathcal{T}_A$, then

$$\mathbf{I}^{\mathfrak{W}}_{\mathbf{w}(0)}(\mathfrak{w}) = \mathbf{I}^{\mathfrak{A}}\left(\mathcal{T}^{-1}(\mathfrak{w})\right), \tag{28}$$

and for every other $\mathfrak{w} \in \mathbb{C}(\Re_+^K)$ we set $\mathbf{I}^{\mathfrak{W}}_{\mathbf{w}(0)}(\mathfrak{w})$ to $+\infty$. ■

Remark: Following [15], [16], [17] it is sufficient to prove the existence of unique solutions to (25) using results from [27] for $\mathfrak{a} \in \mathbb{AC}_{\mathbf{0}}$ such that $\mathbf{I}^{\mathfrak{A}}(\mathfrak{a}) < +\infty$ for the LDP result to hold or to an even smaller set of functions that determine the rate function. The characterisation in (27) would also still apply. However, we feel that using the results of [31] provides us with a complete characterisation.

Now that the underlying rate function $\mathbf{I}^{\mathfrak{W}}_{\mathbf{w}(0)}(\cdot)$ has been specified we set out to derive an alternate expression for $\mathbf{J}(\mathbf{x},t)$ that converts the calculus of variations problem to a finite-dimensional optimisation. In the process we will show that for determining the rate function it suffices to consider piece-wise linear functions (illustrated in Figure 1) $\mathfrak{a} \in \mathbb{AC}_{\mathbf{0}}$ determined by two parameters $v \in [0,t]$ for $t \in (0,1]$ and $\boldsymbol{\lambda} \in \Re_+^K \setminus \mathcal{R}(K)$ such that for $u \in [0,1]$ we have

$$\mathbf{a}(u) = \begin{cases} \boldsymbol{\mu} u & \text{if } u \in [0,v]; \\ \boldsymbol{\lambda}(u-v) + \boldsymbol{\mu} v & \text{if } u \in [v,t]; \\ \boldsymbol{\mu}(u-t+v) + \boldsymbol{\lambda}(t-v) & \text{if } u \in [t,1]. \end{cases} \tag{29}$$

This is proved in the following Lemma.

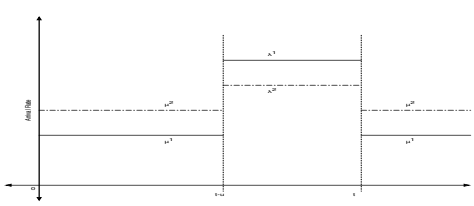

Fig. 1. Typical element of class of $\dot{\mathfrak{a}}$ considered for optimisation.

Lemma 2: If $\chi^{\mathbf{A}}(\mathbf{x}) : \Re_+^K \to \Re_+$ is convex with $\chi^{\mathbf{A}}(\boldsymbol{\mu}) = 0$ for some $\boldsymbol{\mu} \in \Re_+^K$ with $\boldsymbol{\mu} \in \mathcal{R}(K)$ such that there exists $\boldsymbol{\lambda}^* \in \mathcal{R}(K)$ with $\boldsymbol{\mu} < \boldsymbol{\lambda}^*$, then for $\mathbf{x} \in \Re_+^K$ and $t \in (0,1]$ we have

$$\mathbf{J}(\mathbf{x},t) = \begin{cases} \inf_{u \in (0,t]} u \inf_{\boldsymbol{\lambda} \in \arg\max_{\mathbf{r} \in \mathcal{R}} <\mathbf{x},\mathbf{r}>} \chi^{\mathbf{A}}\left(\frac{\mathbf{x}}{u} + \boldsymbol{\lambda}\right) & \text{if } \mathbf{x} \neq \mathbf{0}; \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

Proof: Refer to [18]. ■

Remarks:

1) The workload trajectories (up to time t) that result from the considered class of arrivals is illustrated in Figure 2 where for the sake of illustration, for $v \in [0,t]$ we take $\mathbf{a}(v) = \boldsymbol{\mu}\min(v, t-u) + \boldsymbol{\lambda}(v-t+u)_+$ for some $\boldsymbol{\lambda} \in \Re_+^K \setminus \mathcal{R}(K)$.

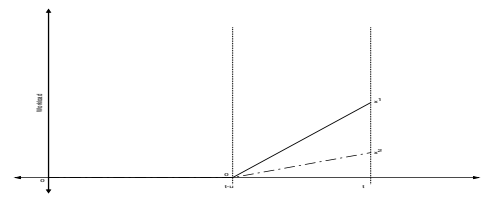

Fig. 2. Typical workload trajectory until time t for the analysed class of input functions $\mathbf{a}$.

2) Lemma 2 informs us that for a given $\mathcal{R}(K)$ to quantify the performance one needs to characterise $H(\mathbf{x})$ for all $\mathbf{x} \in \Re_+^K$. Note that this is equivalent to a complete characterisation of the boundary of $\mathcal{R}(K)$.
3) It is also worth noting that $\mathbf{J}(\mathbf{x},t)$ is convex in its arguments.

Using the expression for $\mathbf{J}(\mathbf{x},t)$ we can now write down simpler expressions for (20), (21), (22) and (23). We have one of the terms of (20) given by

$$\inf_{\mathbf{y} \in \Re_+^K : y^1 \geq x} \inf_{t \in (0,1]} \mathbf{J}(\mathbf{y},t) = x \inf_{z \geq x} \frac{\inf_{\mathbf{y} \in \Re_+^K : y^1 \geq 1} \inf_{\boldsymbol{\lambda} \in H(\mathbf{y})} \chi^{\mathbf{A}}(\mathbf{y}z + \boldsymbol{\lambda})}{z}.$$

We have one of the terms of (21) given by

$$\inf_{\mathbf{y} \in \Re_+^K : y^1 > x} \inf_{t \in (0,1]} \mathbf{J}(\mathbf{y},t) = x \inf_{z \geq x} \frac{\inf_{\mathbf{y} \in \Re_+^K : y^1 > 1} \inf_{\boldsymbol{\lambda} \in H(\mathbf{y})} \chi^{\mathbf{A}}(\mathbf{y}z + \boldsymbol{\lambda})}{z}.$$

Similar logic applied to (22) and (23) yields

$$\inf_{\mathbf{y} \in \Re_+^K : \sum_{k=1}^K y^k \geq x} \inf_{t \in (0,1]} \mathbf{J}(\mathbf{y},t) = x \inf_{z \geq x} \frac{\inf_{\mathbf{y} \in \Re_+^K : \sum_{k=1}^K y^k \geq 1} \inf_{\boldsymbol{\lambda} \in H(\mathbf{y})} \chi^{\mathbf{A}}(\mathbf{y}z + \boldsymbol{\lambda})}{z}; \text{ and}$$

$$\inf_{\mathbf{y} \in \Re_+^K : \sum_{k=1}^K y^k > x} \inf_{t \in (0,1]} \mathbf{J}(\mathbf{y},t) = x \inf_{z \geq x} \frac{\inf_{\mathbf{y} \in \Re_+^K : \sum_{k=1}^K y^k > 1} \inf_{\boldsymbol{\lambda} \in H(\mathbf{y})} \chi^{\mathbf{A}}(\mathbf{y}z + \boldsymbol{\lambda})}{z}.$$

These expressions can be simplified further with additional assumptions on the arrival processes (such as independence) and the rate-region $\mathcal{R}(K)$.

V. Examples

We will present three example rate-regions to show how the analysis developed above applies. The first is a two-user elliptical rate-region. The last two examples are obtained from information theory [1]. The first of these considers a two-user Gaussian broadcast channel and the second a symmetrical two-user multiple-access channel. For the remainder of this section we will set the scheduling weight vector $\boldsymbol{\beta}$ to $(1/2, 1/2)$. Note from the analysis in Section IV that other values of $\boldsymbol{\beta}$ can be analysed by modifying the parameters of $\mathcal{R}(2)$.

A. Example I: A Two-User Queue With An Elliptical Rate-Region

Consider a specific enunciation of our model with two users such that the rate region $\mathcal{R}(2)$ is a quadrant of an ellipse with parameters $r^M, r^m > 0$, i.e.,

$$\mathcal{R}(2) = \left\{ (r^1, r^2) \in \Re_+^2 : \left(\frac{r^1}{r^M}\right)^2 + \left(\frac{r^2}{r^m}\right)^2 \leq 1 \right\}. \quad (31)$$

Now let us solve the scheduling policy generation problem, namely, $H(\mathbf{x}) = \arg\max_{\mathbf{r}\in\mathcal{R}(2)} < \mathbf{x}, \mathbf{r} >$ for $\mathbf{x} \in \Re_+^2$. We will only consider the case of at least one coordinate being positive because $H(\mathbf{0}) = \mathcal{R}(2)$. The optimal solution is the unique point $(\tilde{r}^1, \tilde{r}^2) \in \tilde{\mathcal{R}}(2)$ that satisfies

$$\begin{aligned} \frac{\tilde{r}^1}{r^M} &= \frac{r^M x^1}{\sqrt{(r^M x^1)^2 + (r^m x^2)^2}} \text{ and} \\ \frac{\tilde{r}^2}{r^m} &= \frac{r^m x^2}{\sqrt{(r^M x^1)^2 + (r^m x^2)^2}}, \end{aligned} \quad (32)$$

This can be derived easily using Lagrange multipliers. The rate-region and the solution of the scheduling problem are illustrated in Figure 3.

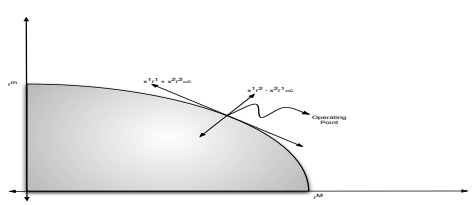

Fig. 3. Illustration of an elliptical rate-region with the solution of the scheduling problem shown.

The expression for $\mathbf{J}(\mathbf{x}, t)$ with $\mathbf{x} \neq \mathbf{0}$ for this example is given by

$$\begin{aligned} &\mathbf{J}(\mathbf{x}, t) = \\ &\inf_{u\in(0,t]} u\, \chi^{\mathbf{A}} \left(\left(\frac{x^1}{u} + r^M \frac{r^M x^1}{\sqrt{(r^M x^1)^2 + (r^m x^2)^2}}, \right.\right. \\ &\qquad \left.\left. \frac{x^2}{u} + r^m \frac{r^m x^2}{\sqrt{(r^M x^1)^2 + (r^m x^2)^2}} \right) \right). \end{aligned} \quad (33)$$

B. Example II: Two-user Gaussian Broadcast Channel

The broadcast channel [1, Section 14.6] models a communication system where there is one transmitter and multiple receivers who can all listen to the transmitter. The capacity region (in natural units, i.e., nats) of a two user Gaussian broadcast channel [1, Section 14.6] is determined by two parameters (signal to noise ratios) $P_1 > P_2 > 0$ and is given as follows:

$$\begin{aligned} \mathcal{R}(2) = \bigcup_{\gamma\in[0,1]} \Bigg\{ &(r^1, r^2) \in \Re_+ 2 : r^1 \leq \frac{1}{2}\log\left(1 + \gamma P_1\right), \\ &r^2 \leq \frac{1}{2}\log\left(\frac{1+P_2}{1+\gamma P_2}\right) \Bigg\}. \end{aligned} \quad (34)$$

If $P_1 = P_2 > 0$, then one gets a simplex.

The scheduling rule $H(x^1, x^2)$ with at least one coordinate positive is then given by $\left(\frac{1}{2}\log(1+\gamma^* P_1), \frac{1}{2}\log\left(\frac{1+P_2}{1+\gamma^* P_2}\right)\right)$ with

$$\gamma^* = \begin{cases} 1 & \text{if } x^1\left(1+\frac{1}{P_2}\right) \geq x^2\left(1+\frac{1}{P_1}\right); \\ 0 & \text{if } x^1 P_1 \leq x^2 P_2; \\ \frac{\frac{x^1}{P_2} - \frac{x^2}{P_1}}{x^2 - x^1} & \text{otherwise.} \end{cases} \quad (35)$$

Again $H(0,0) = \mathcal{R}(2)$.

From this exercise we can now write down expressions for $\mathbf{J}(\mathbf{x}, t)$ for $\mathbf{x} \neq \mathbf{0}$ as follows. If $x^1\left(1+\frac{1}{P_2}\right) \geq x^2\left(1+\frac{1}{P_1}\right)$, then

$$\mathbf{J}(\mathbf{x}, t) = \inf_{u\in(0,t]} u\, \chi^{\mathbf{A}} \left(\frac{x^1}{u} + \frac{1}{2}\log(1+P_1), \frac{x^2}{u} \right); \quad (36)$$

if $x^1 P_1 \leq x^2 P_2$, then

$$\mathbf{J}(\mathbf{x}, t) = \inf_{u\in(0,t]} u\, \chi^{\mathbf{A}} \left(\frac{x^1}{u}, \frac{x^2}{u} + \frac{1}{2}\log(1+P_2) \right); \quad (37)$$

and if neither of the above two conditions hold, then

$$\begin{aligned} &\mathbf{J}(\mathbf{x}, t) = \\ &\inf_{u\in(0,t]} u\, \chi^{\mathbf{A}} \left(\begin{array}{c} \frac{x^1}{u} + \frac{1}{2}\log\left(\frac{x^1(P_1-P_2)}{P_2(x^2-x^1)}\right), \\ \frac{x^2}{u} + \frac{1}{2}\log\left(\frac{(1+P_2)P_1(x^2-x^1)}{x^2(P_1-P_2)}\right) \end{array} \right). \end{aligned} \quad (38)$$

C. Example III: Centralised Multiple Access Channel

Consider the rate region $\mathcal{R}(K)$ to be the capacity region of a K-user multiple-access channel [1, Section 14.3] and [32]. At the beginning of every transmission interval each of the users communicate their queue-lengths to a centralised scheduler that then determines the operating point to be used. It was shown in [32, Lemma 3.4] that $\mathcal{R}(K)$ is a polymatroid [5, Section 11.1] where the rank-function is given by conditional mutual information terms. This rate region is also applicable [33] in the asymptotic regime of (very) high signal-to-noise ratio of the multiple-input, multiple-output multiple-access channel with fading such that the receiver has perfect channel information and the transmitters have no channel information. Such a model [33] exhibits a nice trade-off between diversity and multiplexing that was used to

provide performance bounds based upon the tails of the queue-lengths for max-weight type scheduling algorithms in [34]. In [34] the two-user case was analysed completely when the rate-region is a simplex, and bounds were presented when the rate-region is a symmetric polymatroid. The analysis presented here can be used to improve upon the bounds of [34] in the general case.

Define $\mathcal{K} = \{1, 2, \ldots, K\}$ and suppose that we are given a function $f : \mathfrak{P}(\mathcal{K}) \to \Re_+$ from the power set of $\mathcal{K}$ to the (non-negative) real line. Then the polytope

$$\mathcal{R}_f(K) := \left\{ \mathbf{r} \in \Re_+^K : \sum_{i \in \mathcal{J}} r^i \leq f(\mathcal{J}),\ \mathcal{J} \subseteq \mathcal{K} \right\} \tag{39}$$

is a polymatroid if the function f satisfies the following properties:

1) (*normalised*) $f(\emptyset) = 0$;
2) (*increasing*) if $\mathcal{J}_1 \subseteq \mathcal{J}_2 \subseteq \mathcal{K}$, then $f(\mathcal{J}_1) \leq f(\mathcal{J}_2)$; and
3) (*submodular*) if $\mathcal{J}_1, \mathcal{J}_2 \subseteq \mathcal{K}$, then $f(\mathcal{J}_1) + f(\mathcal{J}_2) \geq f(\mathcal{J}_1 \cup \mathcal{J}_2) + f(\mathcal{J}_1 \cap \mathcal{J}_2)$.

A function f with these properties is called a rank function. Let π be a permutation of $\mathcal{K}$, then the vector $\mathbf{r}_\pi$ defined by

$$\begin{aligned} r_\pi^{\pi(1)} &= f(\{\pi(1)\}) \\ r_\pi^{\pi(2)} &= f(\{\pi(1), \pi(2)\}) - f(\{\pi(1)\}) \\ &\vdots \\ r_\pi^{\pi(K)} &= f(\{\pi(1), \ldots, \pi(K)\}) - f(\{\pi(1), \ldots, \pi(K-1)\}) \end{aligned}$$

belongs to $\mathcal{R}_f(K)$ for all permutations π. Along with $\mathbf{0}$ the points $\mathbf{r}_\pi$ are the extreme points of $\mathcal{R}_f(K)$. Also for any pair of sets $\mathcal{J}_1 \subset \mathcal{J}_2 \subseteq \mathcal{K}$, there exists a point $\mathbf{r} \in \mathcal{R}_f(K)$ such that

$$\sum_{i \in \mathcal{J}_1} r^i = f(\mathcal{J}_1) \text{ and } \sum_{i \in \mathcal{J}_2} r^i = f(\mathcal{J}_2).$$

Maximising a linear functional ($< \mathbf{x}, \mathbf{r} >$) over a polymatroid is very easy [5, Section 11.1.2] and is given by the following:

1) without loss of generality assume that $x^k \geq 0$ for all $k \in \mathcal{K}$. Otherwise we simply set the corresponding $r^k = 0$ for the optimal solution;
2) let π be a permutation of $\mathcal{K}$ such that the weights are in decreasing order, i.e.,

$$x_{\pi(1)} \geq x_{\pi(2)} \geq \cdots \geq x_{\pi(K)} \geq 0, \tag{40}$$

   then $\mathbf{r}_\pi$ is an optimal solution; and
3) the set of optimal solutions is the convex hull of $\mathbf{r}_\pi$ for all permutations of $\mathcal{K}$ that yield the ordering in (40).

With $\mathcal{R}_f(K)$ as the rate-region for our system, this completely specifies $H(\mathbf{x})$ for all $\mathbf{x} \in \Re_+^K$. To understand this better we will look at a class of two-user channels.

1) Symmetric Two-User Case: Let $K = 2$ and given two parameters $r^M > r^m > 0$ define the rank function f as follows:

$$f(\mathcal{J}) := \begin{cases} 0 & \text{if } \mathcal{J} = \emptyset; \\ r^M & \text{if } \mathcal{J} = \{1\}; \\ r^M & \text{if } \mathcal{J} = \{2\}; \\ r^M + r^m & \text{if } \mathcal{J} = \{1, 2\}. \end{cases} \tag{41}$$

Then our rate region is $\mathcal{R}_f(2)$. The edge cases for the parameters r^M, r^m do not give any new insights: if $r^M = r^m > 0$, then $\mathcal{R}_f(2)$ is a square; and if $r^M > r^m = 0$, then $\mathcal{R}_f(2)$ is a simplex. An example of this rate-region is show in Figure 4.

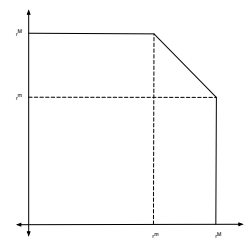

Fig. 4. An example of a symmetrical two-user polymatroidal rate-region.

We now solve for $H(\mathbf{x}) = \arg\max_{\mathbf{r} \in \mathcal{R}_f(2)} < \mathbf{x}, \mathbf{r} >$ for $\mathbf{w} \in \Re_+^2$. We need to partition $\Re_+^2$ into 6 regions; these and the corresponding sets $H(\mathbf{x})$ are:

1) Region $A = \{\mathbf{0}\}$. Here it is clear that $H(\mathbf{0}) = \mathcal{R}(2)$.
2) Region $B = \{x^1 > 0,\ x^2 = 0\}$. Then $H(\mathbf{x}) = \{\tilde{r}^1 = r^M,\ \tilde{r}^2 \in [0, r^m]\}$.
3) Region $C = \{x^1 > x^2 > 0\}$. Then $H(\mathbf{x}) = \{\tilde{r}^1 = r^M,\ \tilde{r}^2 = r^m\}$.
4) Region $D = \{x^1 = x^2 > 0\}$. Then $H(\mathbf{x}) = \{(\tilde{r}^1, \tilde{r}^2) \in [r^m, r^M]^2 :\ \tilde{r}^1 + \tilde{r}^2 = r^M + r^m\}$.
5) Region $E = \{x^2 > x^1 > 0\}$. Then $H(\mathbf{x}) = \{\tilde{r}^1 = r^m,\ \tilde{r}^2 = r^M\}$.
6) Region $F = \{x^2 > 0,\ x^1 = 0\}$. Then $H(\mathbf{x}) = \{\tilde{r}^1 \in [0, r^m],\ \tilde{r}^2 = r^M\}$.

The different scheduling regions are shown in Figure 5.

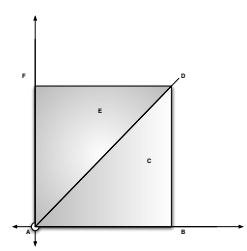

Fig. 5. Partitioning of $\Re_+^2$ into regions where the same scheduling action results.

Using the above we can write down expressions for $\mathbf{J}(\mathbf{x}, t)$ as follows:

1) if $\mathbf{x} \in B$, then

$$\mathbf{J}(\mathbf{x},t) = \inf_{u\in(0,t]} u \inf_{r^2\in[0,r^m]} \chi^{\mathbf{A}}\left(\frac{x^1}{u}+r^M, r^2\right); \quad (42)$$

2) if $\mathbf{x} \in C$, then

$$\mathbf{J}(\mathbf{x},t) = \inf_{u\in(0,t]} u\, \chi^{\mathbf{A}}\left(\frac{x^1}{u}+r^M, \frac{x^2}{u}+r^m\right); \quad (43)$$

3) if $\mathbf{x} \in D$, then

$$\mathbf{J}(\mathbf{x},t) = \inf_{u\in(0,t]} u \inf_{\substack{(r^1,r^2)\in[r^m,r^M]^2:\\ r^1+r^2=r^M+r^m}} \chi^{\mathbf{A}}\left(\frac{x^1}{u}+r^1, \frac{x^2}{u}+r^2\right); \quad (44)$$

4) if $\mathbf{x} \in E$, then

$$\mathbf{J}(\mathbf{x},t) = \inf_{u\in(0,t]} u\, \chi^{\mathbf{A}}\left(\frac{x^1}{u}+r^m, \frac{x^2}{u}+r^M\right); \quad (45)$$

5) if $\mathbf{x} \in F$, then

$$\mathbf{J}(\mathbf{x},t) = \inf_{u\in(0,t]} u \inf_{r^1\in[0,r^m]} \chi^{\mathbf{A}}\left(r^1, \frac{x^2}{u}+r^M\right). \quad (46)$$

ACKNOWLEDGEMENTS

This work was supported by SFI grant 03/IN3/I396. The author would like to thank Ken Duffy and Tolya Puhalskii for detailed suggestions and comments that were critical in expanding the scope and improving the presentation of this paper.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd ed. Hoboken, NJ: Wiley-Interscience [John Wiley & Sons], 2006.

[2] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1936–1948, 1992.

[3] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Trans. on Comm.*, vol. 47, no. 8, pp. 1260–1267, Aug. 1999.

[4] M. Andrews, K. K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queueing system with asynchronously varying service rates," *Probability in Engineering and Informational Sciences*, vol. 18, no. 2, pp. 191–217, 2004.

[5] H. Chen and D. D. Yao, *Fundamentals of Queueing Networks*. Springer, New York, 2001.

[6] S. Foss and T. Konstantopoulos, "An overview of some stochastic stability methods," *Journal of the Operations Research Society of Japan*, vol. 47, no. 4, pp. 275–303, 2004.

[7] C. Maglaras and J. A. Van Mieghem, "Queueing systems with leadtime constraints: A fluid-model approach for admission and sequencing control," *European J. Oper. Res.*, vol. 167, no. 1, pp. 179–207, 2005.

[8] A. L. Stolyar and K. Ramanan, "Largest weighted delay first scheduling: Large deviations and optimality," *The Annals of Applied Probability*, vol. 11, no. 1, pp. 1–48, 2001.

[9] A. L. Stolyar, "Control of end-to-end delay tails in a multiclass network: LWDF discipline optimality," *Ann. Appl. Probab.*, vol. 13, no. 3, pp. 1151–1206, 2003.

[10] D. Bertsimas, I. Paschalidis, and J. Tsitsiklis, "Asymptotic buffer overflow probabilities in multi-class multiplexers: An optimal control approach," *IEEE Trans. on Automatic Control*, vol. 43, pp. 315–335, March 1998.

[11] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications, 2nd Ed.* Springer, 1998.

[12] P. Billingsley, *Convergence of probability measures*, 2nd ed., ser. Wiley Series in Probability and Statistics: Probability and Statistics. New York: John Wiley & Sons Inc., 1999, a Wiley-Interscience Publication.

[13] S. N. Ethier and T. G. Kurtz, *Markov processes: Characterization and convergence*, ser. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons Inc., 1986.

[14] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, 2nd ed., ser. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag, 2003, vol. 288.

[15] A. A. Puhalskii, *Large deviations and idempotent probability*, ser. Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics. Chapman & Hall/CRC, Boca Raton, FL, 2001, vol. 119.

[16] ——, "The action functional for the Jackson network," *Markov Processes and Related Fields*, vol. 13, no. 1, pp. 99–136, 2007.

[17] A. A. Puhalskii and A. A. Vladimirov, "A large deviation principle for Join the Shortest Queue," *Mathematics of Operations Research*, vol. 32, pp. 700–710, 2007.

[18] V. G. Subramanian, "Large deviations of max-weight scheduling policies on convex rate regions," Hamilton Institute, National University of Ireland, Maynooth, Co. Kildare, Ireland, Tech. Rep., 2007.

[19] J. Garcia, "An extension of the contraction principle," *J. Theoret. Probab.*, vol. 17, no. 2, pp. 403–434, 2004.

[20] J. L. Doob, *Measure theory*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 1994, vol. 143.

[21] A. Dembo and T. Zajic, "Large deviations: From empirical mean and measure to partial sums process," *Stochastic Process. Appl.*, vol. 57, no. 2, pp. 191–224, 1995.

[22] R. M. Dudley, *Real analysis and probability*, ser. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 2002, vol. 74.

[23] A. A. Puhalskii, "Large deviations for stochastic processes," Notes for LMS/EPSRC Short Course: Stochastic Stability, Large Deviations and Coupling Methods, Heriot-Watt University, Edinburgh, Sept. 2006. [Online]. Available: http://www-math.cudenver.edu/~puhalski/ld/course_1.pdf

[24] A. A. Mogul′skiĭ, "Large deviations for the trajectories of multidimensional random walks," *Theory Probab. Appl.*, vol. 21, no. 2, pp. 309–323, 1976.

[25] ——, "Large deviations for processes with independent increments," *Ann. Probab.*, vol. 21, no. 1, pp. 202–215, 1993.

[26] K. Yosida, *Functional analysis*, reprint of the sixth (1980) ed., ser. Classics in Mathematics. Berlin: Springer-Verlag, 1995.

[27] H. Brézis, *Opérateurs maximaux monotones et semigroups de contractions dans les espaces de Hilbert*. Amsterdam: North-Holland Publishing Co., 1973, north-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[28] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, ser. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag, 1998, vol. 317.

[29] F. E. Browder, "Nonlinear operators and nonlinear equations of evolution in Banach spaces," in *Nonlinear functional analysis (Proc. Sympos. Pure Math., Vol. XVIII, Part 2, Chicago, Ill., 1968)*. Providence, R.I.: Amer. Math. Soc., 1976, pp. 1–308.

[30] R. Rockafellar, *Convex analysis*, ser. Princeton Mathematical Series, No. 28. Princeton, N.J.: Princeton University Press, 1970.

[31] E. Cépa, "Problème de Skorohod multivoque," *Ann. Probab.*, vol. 26, no. 2, pp. 500–532, 1998.

[32] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, 1998.

[33] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Information Theory*, vol. 50, no. 9, pp. 1859–1874, 2004.

[34] S. Kittipiyakul and T. Javidi, "Optimal operating point for MIMO multiple access channel with bursty traffic," *accepted to IEEE Transactions on Wireless Communication*, 2007.