On the Capacity of Information Processing Systems

Laurent Massoulié
Microsoft Research-Inria
Joint Centre
laurent.massoulie@inria.fr

Kuang Xu
Graduate School of Business
Stanford University
kuangxu@stanford.edu

Abstract

We propose and analyze a family of *information processing systems*, where a finite set of experts or servers are employed to extract information about a stream of incoming jobs. Each job is associated with a hidden label drawn from some prior distribution. An inspection by an expert produces a noisy outcome that depends both on the job's hidden label and the type of the expert, and occupies the expert for a finite time duration. A decision maker's task is to dynamically assign inspections so that the resulting outcomes can be used to accurately recover the labels of all jobs, while keeping the system stable. Among our chief motivations are applications in crowd-sourcing, diagnostics, and experiment designs, where one wishes to efficiently learn the nature of a large number of items, using a finite pool of computational resources or human agents.

We focus on the *capacity* of such an information processing system. Given a level of accuracy guarantee, we ask how many experts are needed in order to stabilize the system, and through what inspection architecture. Our main result provides an adaptive inspection policy that is asymptotically optimal in the following sense: the ratio between the required number of experts under our policy and the theoretical optimal converges to one, as the probability of error in label recovery, δ, tends to zero.[1]

keywords: stochastic resource allocation, hypothesis testing, information processing system, fluid model.

1 Introduction

An increasing number of modern processing systems have been designed and deployed for the purpose of learning and information extraction. In these applications, which we refer to broadly as **information processing systems**, a group of experts or servers is tasked with performing (noisy) inspections on a large collection of jobs, with the objective of uncovering some hidden features associated with each job up to a level of desirable accuracy. Below are some examples:

1. *Crowd-sourcing* (Karger et al. 2014, Ho et al. 2013, Bernstein et al. 2012): a collection of images is dispatched to a group of human agents, where an agent attaches a label to each assigned image based on her own judgment. A decision maker then aggregates agents' responses to produce a "best" label for each image.
2. *Medical diagnostics* (Gerdtz and Bucknall (2001)): medical data of patients is reviewed by physicians or nurses with different domains of expertise, with the goal of correctly identifying the patients' diseases.
3. *Quality management* (Baker and von Beers (1996)): a set of products undergoes a number of different tests performed by specialized machines, to identify whether a product is faulty and the type of fault it contains.

[1]March 2016; revised February 2017. An extended abstract of this paper was presented at the Conference on Learning Theory (COLT) 2016. The authors would like to thank the anonymous referees at Operations Research and COLT for their detailed, constructive feedback.

The presence of **resource constraints** is a crucial feature shared across many of these systems: the amount of processing resources, such as human agents, machines, or computer servers, is finite, and yet, each inspection or test requires the corresponding resource to commit a non-trivial amount of effort. This raises a natural question:

How much information can we extract using a finite amount of processing resources?

The main objective of the present paper is to address this question, and gain understanding of the "capacity" of an information processing system. We will approach this problem by studying the minimum required system size, defined, roughly speaking, as the minimum number of experts needed in order to learn a sufficient amount of information about *every* job in a stream of arrivals, while ensuring system stability. [2]

We now informally describe our model. The system consists of m experts with different *types* (expertise), where the fraction of experts of type k is r_k. The system receives a stream of incoming jobs arriving at unit rate, where each job is associated with a *label*, hidden from the decision maker, which is drawn i.i.d. from some prior distribution, π. An atomic unit of processing is called an *inspection*: the decision maker may assign an expert to perform an inspection on a job, which occupies the expert for a random period of time, with a mean that depends on the type of the expert. We will use the terms *initiation* and *completion* to denote the start and finish of an inspection, respectively. The inspection leads to a (noisy) *outcome*, whose distribution, $p(k, h, \cdot)$, depends on both the type of the expert, k, and the true label of the job under inspection, h. The goal of the decision maker is to assign inspections intelligently, and use the resulting outcomes to produce, for each job, a *classification* (i.e., prediction) of its hidden label. We say that the system is *s*table if all jobs will receive a classification in a finite amount of time.

Note that we cannot have a meaningful discussion on the resource requirement of this system without specifying how accurate the classifications need to be. Indeed, in the absence of any accuracy constraint, the decision maker can simply make up classifications without performing a single inspection, and the system would always be trivially stable. Therefore, we will designate an accuracy parameter, $\delta \in (0, 1)$, and demand that the probability of mis-classification for each job be at most δ. Since inspections take up the experts' bandwidth, we expect that a smaller δ would demand more inspections, which, in turn, require more experts.

The main goal of this paper is to understand the *minimum number of experts* needed for a given accuracy parameter, δ, and how to achieve it via intelligent policies. This definition of minimum system size is motivated by the intuition that a decision policy that requires the least number of experts to stabilize the system also, in a sense, most efficiently utilizes the processing resources.

Preview of main result. The main result of this paper proposes an inspection architecture which, for *any* non-trivial outcome distributions, asymptotically achieves the minimum system size in the regime of high accuracy ($\delta \to 0$). Specifically, the ratio between the required number of experts under our policy, m, and that of a theoretical minimum, m^*, converges to 1 uniformly across all prior distributions of job labels, as $\delta \to 0$:

$$\sup_{\pi} \frac{m}{m^*} = 1 + \mathcal{O}\left(\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}\right), \quad \text{as } \delta \to 0. \tag{1}$$

Moreover, the policy does not require knowing the prior distribution, π, but adapts to it automatically.

We conclude this section by highlighting two main design challenges in creating an efficient decision policy for information processing systems. The first challenge arises from the fact that the processing resources are often heterogeneous, a result of the variations in expertise, machine functionality, or personal trait. In our model, this is captured by the fact that the outcomes distributions, $\{p(h, k, \cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$, may vary significantly depending on the type of the inspecting

[2]Throughout the paper, we will use the term "expert" to refer to a single unit of processing resource, with the understanding that it may represent a computer server, testing machine, or human agent, depending on the application.

expert. A key implication of such heterogeneity is that the decision policy must be sufficiently adaptive to a job's past inspection outcomes, because depending on what we "believe" the job's label to be, some combinations of expert types may be more efficient than others. For instance, consider a testing system where job labels correspond to three domains that might have caused an error in a product: {network, hardware, software}, and the experts are of two kinds:

1. General experts: they know a little about all domains, and produce noisy inspection outcomes.
2. Specialists: who have deep expertise in one of the three domains but know nothing about the other two. Their inspections contribute strong signals towards confirming an error in their own domain, but are otherwise non-informative if the job's true label lies elsewhere.

Intuitively, a sensible decision policy for this setting should be adaptive: a job can be first inspected by a few general experts to zone in on a possible domain, and depending on their opinions, the job will then be sent to the corresponding specialists to further "confirm" the diagnosis, with some small possibility of back-and-forth if the initial diagnosis had been incorrect. In contrast, sending a job directly to a specialist at the beginning would risk the specialist being of the wrong kind, and never sending a job to any specialist would overwhelm the generalists who can only provide limited information per inspection; in both cases, there would be waste of processing resources.

The second challenge is more nuanced, and stems from the combined effect of the resource constraint and expert heterogeneity: the "optimal" course of inspections for an individual job does not only depend on the types of experts available, but also on the *prior distribution* governing the proportions of labels among its fellow jobs. In other words, there will be *contention* among jobs as they "compete" for the same processing resources. For instance, in the above-mentioned testing example, suppose there emerges a disproportionately large fraction of jobs with the label "network", while the fraction of specialists in "network" remains fixed. In such a case, it is plausible that some of the generalists may now need to be enlisted to provide assistance by performing more inspections on these jobs than they would before. Therefore, the mixture of inspections that a job receives may shift as the prior distribution changes, which shows that a good decision policy cannot be overly centered around individual jobs, and must be aware of the overall arrival pattern.

1.1 Related Research

The present paper intersects with two main areas of research: statistics and dynamic resource allocation. Most related to our work is the literature on sequential hypothesis testing in statistics, dating back to the seminal work of Wald (1945), which studies the problem of distinguishing hypotheses of a distribution by sequentially drawing samples from it, with the objective of minimizing a combined cost involving the number of samples and the resulting probability of error (see also Siegmund (2013) and the references therein). Notably, Chernoff (1959) considers an important generalization of Wald's problem, where, instead of drawing samples from the same distribution and deciding when to stop, the decision maker has the additional freedom to choose from *multiple* available experiments, and the distribution of the experimental outcome depends on both the true hypothesis and the type of the experiment. Chernoff identifies a dynamic testing policy that asymptotically achieves the minimum sample complexity, and computes explicitly the leading factor via the solution to a zero-sum game. The current paper draws inspiration from this literature, and especially the multi-experiment version of Chernoff (1959). Yet, there are some key differences: while sequential hypothesis testing aims to reduce the sample complexity associated with testing a *single* distribution, we are interested in performing tests for *multiple* jobs simultaneously using finite resources. The resource constraint creates contention and coupling among the jobs, and as was alluded to in Introduction, policies designed to minimize sample complexity for testing a single hypothesis do not easily extend to our problem.

Our work is also related to the literature on dynamic resource allocation, and particularly multi-class, multi-server queueing networks (Harrison and López (1999), Talreja and Whitt

(2008), Tsitsiklis and Xu (2017)). In our model the inspection outcomes may have different distributions depending on the expert's type and the job's label, which is roughly analogous to a queueing network where the service rate depends on the class of the server and the job being served. However, our system differs from this literature in two major aspects. First, while in conventional processing systems jobs typically come with an *exogenous* service requirement, also referred to as workload or job size, in our system a job's service requirement is defined *endogenously* in relation to how much information needs to be gathered in order to uncover its label. Second, in a queueing network the class of an incoming job is typically known to the decision maker, while in our model the job labels are hidden, and the very objective of processing is to uncover them.

There have also been a growing interest in the connections between information acquisition and resource allocation (Alizamir et al. (2013), Bimpikis and Markakis (2016), Johari et al. (2016), Harrison and Sunar (2015), Retsef et al. (2015)). They are similar to our work in spirit, but differ in models and objectives. The authors of Alizamir et al. (2013) study a single-server queue where the server decides on how many tests to perform on each job to reach a binary diagnosis, while achieving an optimal delay-accuracy trade-off. The information structure in Alizamir et al. (2013) is substantially more restricted than in our model: the tests are identical, and outcomes and job types are binary; on the other hand, Alizamir et al. (2013) analyzes queueing delay which we do not consider. The recent papers Johari et al. (2016) and Bimpikis and Markakis (2016) study processing systems with heterogeneous job and server types and capacitated processing resources, where, similar to our model, it is important for the decision maker to learn about the job types in order to identify the best processing scheme. These models differ from ours in that learning serves as a means towards another objective, such as maximizing total rewards in Johari et al. (2016) or improving throughput or delay in Bimpikis and Markakis (2016) and Retsef et al. (2015), whereas information extraction is the intrinsic purpose of processing in our problem.

On the technical end, we build on several existing techniques and ideas. To establish the stability of our decision policy, which involves three interconnected stages, we will leverage a program pioneered by Rybko and Stolyar (1992) and Dai (1995) in the context of queueing networks, which approximates the system dynamics using a certain fluid limit, and subsequently uses a contraction property of the fluid limits to derive the stability of the original system. A subroutine of our inspection policy dynamically creates workload vectors by dynamically solving a linear program to minimize the incremental change to a potential function, which is reminiscent of, and inspired by, the family of max-weight scheduling policies (Tassiulas and Ephremides 1992). Finally, to derive upper and lower bounds on the probability of error under an inspection policy, we make elementary uses of well-known techniques in probability theory and statistics, such as changes of measures and concentration inequalities for martingales.

2 Model and Metrics

2.1 The Model

System primitives. The system evolves in continuous time, indexed by $t \in \mathbb{R}_+$. There is a stream of jobs that arrives to the system according to a Poisson process with rate $\lambda_0 > 0$. Without loss of generality, we assume that $\lambda_0 = 1$, because one can simply scale the expressions of system size by λ_0 so that the results in the paper apply to other values of λ_0 as well. We index jobs by the order in which they arrive, and refer to the ith job that arrives to the system as *job i*. Job i is associated with a *label*, H_i, which belongs to a finite set, $\mathcal{H}$, whose cardinality is $c_{\mathcal{H}}$. The job labels are independent and identically distributed according to a prior distribution, $\pi = \{\pi_h\}_{h\in\mathcal{H}}$, with $\mathbb{P}(H_1 = h) = \pi_h$, and the labels are unknown to the decision maker. We assume all entries of π to be positive.

The system is equipped with m *experts*. Let $\mathcal{E}$ be the set of experts. Each expert is associated with a *type*, k, from a finite set, $\mathcal{K}$, with cardinality $c_{\mathcal{K}}$. The number of type-k experts is $\rho_k m$, $k \in \mathcal{K}$, where ρ_k is the fraction of type-k experts in the system, with $\rho_k > 0$ and $\sum_{k\in\mathcal{K}} \rho_k = 1$.

We will refer to $\{\rho_k\}_{k\in\mathcal{K}}$ as the *expert mixture*.[3]

Inspections and resource constraints. An expert can be called upon to perform an inspection of any job in the system. At the end of an inspection, a random *outcome* is produced which takes values in a finite set, $\mathcal{X}$. We denote by $X_{i,j}$ the outcome of the jth inspection performed on job i. Suppose that the true label of job i is h, and that the jth inspection is performed by an expert of type k, then $X_{i,j}$ is a random variable distributed according to the outcome distribution, $p(h,k,\cdot)$, and is independent from all other parts of the system. Note that this implies also that an expert may inspect a job multiple times, producing i.i.d. outcomes[4]. The diversity in the set of outcome distributions, $\{p(h,k,\cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$, captures the possibility that experts of various types may have different expertise. We assume the set of outcome distributions is known to the decision maker.

The pool of experts is *resource constrained*, in the sense that each individual expert can only perform, on average, a finite number of inspections per unit time. Formally, at any time t, we assume that each of the m experts can be in one of the two states: IDLE and BUSY, and all experts are initialized in state IDLE. An IDLE expert of type k can be assigned to initiate an inspection of a job currently in the system. Once the inspection starts, the expert enters the state BUSY for a duration that is exponentially distributed with mean $1/\mu_k$, $\mu_k > 0$, independent from the rest of the system. We refer to $\{\mu_k\}_{k\in\mathcal{K}}$ as the *inspection rates*, since μ_k corresponds to the average number of inspections that a type-k expert can perform in unit time. The expert returns to the IDLE state once the inspection is completed. The inspections are non-preemptive, so that an expert cannot start inspecting a different job before the previous inspection has been finished. We assume that multiple experts are allowed to inspect the same job at the same time. The latter assumption is motivated by applications where jobs, such as images or data files, can be duplicated at relatively low costs or be accessible to multiple experts concurrently.

Note that the inspection rate depends on the expert's type but not the job's hidden label. Similar to the earlier assumption that the arrival rate $\lambda_0 = 1$, without loss of generality, we assume that $\{\mu_k\}_{k\in\mathcal{K}}$ is normalized so that the average inspection rate across different expert types is 1:

$$\overline{\mu} \overset{\triangle}{=} \sum_{k\in\mathcal{K}} \rho_k \mu_k = 1. \tag{2}$$

Departure rule. At any time t, the system operator can choose to let a job depart from the system. Upon job i's departure from the system, the operator must produce a *classification*, $\widehat{H}_i$, representing her belief of job i's true label. We say that there is an *error* if the classification does not match the true label, i.e., if $\widehat{H}_i \neq H_i$.

2.2 Conditions on Outcome Distributions

The set of outcome distributions, $\{p(h,k,\cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$, plays a central role in our model, because the job labels can only be differentiated via the inspection outcomes. In the present paper, we allow for essentially any outcome distribution over $\mathcal{X}$, with the exception of two conditions. Informally, we assume that (1) all outcomes are "noisy", so that no single outcome can distinguish between two job labels with certainty, and (2) for any two job types, there exists at least one type of experts who can distinguish them. Neither assumption leads to severe a loss of generality, as we explain shortly.

We now give formal definitions of the two conditions. In our regime of interest, where the target classification error is small, it turns out that an important measure of the informativeness

[3]To avoid excessive use of floors and ceilings in the notation, we shall assume throughout that the parameters are chosen in such a way that $\rho_k m$ are always positive integers. Since we are mostly interested in the regime where m is large, this assumption does not affect the main result in a substantial way. The main algorithm and analytical results can also be extended should floors or ceilings be necessary.

[4]The i.i.d. assumption is equivalent to having the outcomes solely determined by the expert's type, the job's true label, and some exogenous noise.

of an inspection outcome is that of KL-divergence, defined as follows. Fix $i, j \in \mathbb{N}$, h, l in $\mathcal{H}$, and $k \in \mathcal{K}$, denote by $Z_{i,j}(h, l, k)$ the likelihood associated with the jth inspection to job i done by a type-k expert, as follows:

$$Z_{i,j}(h, l, k) = \ln \frac{p(h, k, X_{i,j})}{p(l, k, X_{i,j})}. \tag{3}$$

The KL-divergence from the outcome distribution $p(h, k, \cdot)$ to $p(l, k, \cdot)$, denoted by $D(h, l, k)$, is defined by the expected value of $Z_{i,j}(h, l, k)$ conditional on the true label of job i being h:

$$D(h, l, k) = \mathbb{E}\left(Z_{1,1}(h, l, k) \,\middle|\, H_1 = h\right) = \sum_{x \in \mathbb{N}} p(h, k, x) \ln \frac{p(h, k, x)}{p(l, k, x)}. \tag{4}$$

Intuitively, the value of $D(h, l, k)$ captures the ability of an expert of type k in telling apart whether a job's label is h versus l; a higher value of $D(h, l, k)$ indicates that the expert's inspection, on average, provides stronger evidence that the true label of a job is more likely to be h than l.

We assume that the outcome distributions $\{p(h, k, \cdot)\}_{h \in \mathcal{H}, k \in \mathcal{K}}$ satisfy two conditions as follows, expressed in terms of KL-divergence.

1. The outcome distributions should be sufficiently diverse so that accurate classifications are possible. For any two distinct job labels, $h, l \in \mathcal{H}$, we assume that there exists at least one expert type, k, for whom the outcome distributions, $p(h, k, \cdot)$ and $p(l, k, \cdot)$, are non-identical. This is equivalent to saying that

$$\min_{h, l \in \mathcal{H}, h \neq l} \max_{k \in \mathcal{K}} D(h, l, k) > 0. \tag{5}$$

   We will define $\underline{d}$ as the smallest non-zero value among the $D(h, l, k)$:

$$\underline{d} = \min_{\substack{h, l \in \mathcal{H}, k \in \mathcal{K}, \\ D(h, l, k) > 0}} D(h, l, k). \tag{6}$$

2. All KL-divergences should be finite, so that a single inspection cannot distinguish two job labels with absolute certainty:

$$\overline{d} = \max_{\substack{h, l \in \mathcal{H},\, h \neq l \\ k \in \mathcal{K}}} D(h, l, k) < \infty. \tag{7}$$

We note that the above two conditions do not significantly restrict the outcome distributions. The first condition is in fact necessary for the problem to be non-trivial, for otherwise one would not be able to distinguish between jobs of labels h and l. The second condition requires all inspections to be "noisy", and it is a natural assumption in applications where the inspections have some intrinsic variability, such as when inspections are performed by human agents, or by machines that are subject to idiosyncratic noise.

2.2.1 Example Outcome Distributions

For concreteness, we describe here a simple example of a family of outcome distributions. Consider the case where a job is an image, containing one out of three possible animals, with $\mathcal{H} = \{\text{cat}, \text{dog}, \text{rabbit}\}$. There are three types of experts, with $\mathcal{K} = \{1, 2, 3\}$, who are asked to perform inspections leading to binary outcomes (e.g., "like" or "dislike"), with $\mathcal{X} = \{0, 1\}$. Fix $p, q \in (0, 1)$, $p \neq q$. Denote by A and B the Bernoulli distribution with mean p and q, respectively. The outcome distributions, $p(h, k, \cdot)$, are given by the following matrix, where the column corresponds to the labels and the row to the expert types: $\begin{pmatrix} A & A & B \\ A & B & A \\ B & A & A \end{pmatrix}$. For instance, the entry $(2, 1)$ indicates that a type-2 expert's inspection of an image containing a 'cat' is distributed according to A. Note that the outcomes are statistically identical when a type-1 expert inspects an image with a cat versus a dog, but are different from the outcome when the image contains a rabbit. However, it is not difficult to see that if we were to obtain many inspections from any two expert types, one could eventually uncover the true label of an image. This example illustrates that for the overall processing system to be effective, it is not necessary that a single expert be able to distinguish all job labels.

2.3 Inspection Policies and Performance Metrics

Inspection policies. To facilitate our discussion, we now introduce the concept of an inspection policy. An inspection policy, ψ, has access to the entire system state and all past history. At any time t, it has the ability to: (1) let an IDLE expert initiate an inspection on a job; (2) let a job i depart from the system, in which case the policy will have to produce a classification, $\widehat{H}_i$, for the job's label, H_i. In addition, the inspection policy can take as input the following parameters: (1) the number of experts, m, the expert mixture, $\{\rho_k\}_{k\in\mathcal{K}}$, and the inspection rates, $\{\mu_k\}_{k\in\mathcal{K}}$; (2) the outcome distributions, $\{p(h,k,\cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$; (3) an accuracy parameter, δ; (4) (potentially) the prior distribution, π, of the job labels. An inspection policy that does not require the knowledge of the prior distribution, π, is said to be *prior-oblivious*.

In our system, there are two main performance criteria for an inspection policy that are of interest. First, we would like an inspection policy to accurately recover the true labels of all jobs, quantified in the following definition.

Definition 2.1 (δ-accuracy) *A policy is δ-**accurate**, if for all $h \in \mathcal{H}$ and $i \in \mathbb{N}$, we have that $\mathbb{P}(\widehat{H}_i \neq h \,|\, H_i = h) \leq \delta$.*

That is, the resulting probability of misclassification on any job of type h is at most δ under a δ-accurate policy.[5] Since the fundamental task of our processing system is to recover job labels reliably, we assume throughout the paper that an inspection policy should always be δ-accurate, where δ is the accuracy requirement chosen by the decision maker.

In addition to accuracy, a second important benchmark is stability. That is, every job should be able to depart from the system in a finite amount of time. We say that an inspection policy is stable if there exists a time-stationary regime for all processes used to characterize the system's evolution. For instance, a sufficient condition for stability would be that by taking into account all suitable state variables, the corresponding state variable process fully describes the system dynamics and constitutes an irreducible positive-recurrent Markov process. Compared to the definition of accuracy, the notion of stability is more delicate, because whether an inspection policy can stabilize a system depends on the *relative* magnitudes between the number of experts, m, and the accuracy requirement, δ: because the inspections are noisy, as δ decreases, each job will require a larger number of inspections to achieve a desired classification accuracy. Since the arrival rate of jobs is assumed to be fixed, this further implies that the number of experts, m, must also grow accordingly.

Therefore, as alluded to in the Introduction, a natural way of assessing how efficient an inspection policy is at stabilizing the system is to measure the minimum system size (i.e., m) needed in order for the system to be stable for a given accuracy target, δ. This inspires the following notion of **resource efficiency**, where we compare the minimum system size required by an inspection policy against that of a theoretical optimal, as follows. Fix an expert mixture, $\{\rho_k\}_{k\in\mathcal{K}}$, inspection rates, $\{\mu_k\}_{k\in\mathcal{K}}$, and outcome distributions, $\{p(h,k,\cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$. For a policy, ψ, define $m_\psi(\delta,\pi)$ as the smallest system size required under ψ in order to ensure stability:

$$m_\psi(\delta,\pi) = \min\{m \in \mathbb{N} : \text{given } \delta \text{ and } \pi, \text{ a system with } m \text{ experts is stable under } \psi\}. \tag{8}$$

Given prior distribution π and $\delta > 0$, we define $m^*(\delta,\pi)$ as the smallest system size for which there exists a stable inspection policy that is δ-accurate. That is, $m^*(\delta,\pi)$ represents the minimal

[5]Note that the notion of δ-accuracy is defined for every job, and hence implies that the long-run average classification error across all jobs is at most δ.

amount of processing resources required to ensure stability under an "optimal" inspection policy. The following definition serves as the main performance metric of this paper.

Definition 2.2 *We say that an inspection policy, ψ, is* **resource efficient***, if*

$$\limsup_{\delta\to 0} \frac{m_\psi(\delta,\pi)}{m^*(\delta,\pi)} = 1, \quad \textit{for all prior distribution, } \pi. \tag{9}$$

We say that ψ is **strongly resource efficient***, if the above convergence occurs uniformly over all prior distributions:*

$$\limsup_{\delta\to 0} \sup_{\pi} \frac{m_\psi(\delta,\pi)}{m^*(\delta,\pi)} = 1. \tag{10}$$

3 Main Result

The main result of the current paper is the following theorem.

Theorem 3.1 *Fix an expert mixture, $\{\rho_k\}_{k\in\mathcal{K}}$, inspection rates, $\{\mu_k\}_{k\in\mathcal{K}}$, and outcome distributions, $\{p(h,k,\cdot)\}_{h\in\mathcal{H},k\in\mathcal{K}}$. There exists a prior-oblivious, strongly resource efficient inspection policy, ψ. In particular, there exist $c_0, \delta_0 > 0$, such that*

$$\sup_{\pi} \frac{m_\psi(\delta,\pi)}{m^*(\delta,\pi)} \le 1 + c_0 \sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}, \quad \forall \delta \in (0,\delta_0). \tag{11}$$

Notably, one such prior-oblivious, strongly resource-efficient inspection policy will be constructed explicitly in Section 5. The policy is based on a three-stage architecture that first generates a *coarse* label estimate for each job, and subsequently uses the majority of the processing resources to *verify* the validity of the coarse estimates, in an adaptive manner. This policy also inspires a simple heuristic algorithm, discussed in Appendix D, which can be much easier to implement in practice.

We highlight two important features of Theorem 3.1. First, the inspection policy is strongly resource efficient, which implies that its performance guarantee in comparison to the theoretical optimal holds *independently* of the prior distribution. Second, the inspection policy is *prior-oblivious* so that it can operate without any knowledge of the prior distribution of the job labels. This feature is especially important for our problem, since knowing the prior distribution would have likely required first learning the labels of the incoming jobs, which is the very task that we are trying to solve! Moreover, because a prior-oblivious policy automatically adapts to any prior distribution, it is also more robust if the prior distribution were to shift over time, a likely scenario for many applications.

4 Proof Overview and Preliminaries

4.1 The Main Ideas

The remainder of the paper is devoted to the proof of Theorem 3.1. Before delving into the details, we will start by illustrating the main ideas of the proof. Our main goal is to design an inspection architecture to extract information efficiently using a finite number of experts. We will break this general problem further into three sub-problems, in the following order:

(a) What type of information is sufficient for making accurate classification decisions?

(b) How much information do we need to gather for an *individual* job, via inspections, in order to produce an accurate classification of its label?

(c) How can we gather such information for *all* jobs simultaneously in a resource-efficient manner?

We now address each of the three points in order. For point (a), the following notion of cumulative log-likelihood ratio, a concept widely used in statistics, will be central in quantifying information in our problem. Fix $i \in \mathbb{N}$ and $t \in \mathbb{R}_+$. Denote by $N_{i,t}$ the total number of inspections received by job i by time t, and by $K_{i,j}$ the type of the expert who performed the jth inspection on job i. For h and l in $\mathcal{H}$, we define the cumulative log-likelihood ratios for job i at time t as

$$S_{i,t}(h,l) = \sum_{j=1}^{N_{i,t}} Z_{i,j}(h,l,K_{i,j}) = \ln\left(\prod_{j=1}^{N_{i,t}} \frac{p(h,K_{i,j},X_{i,j})}{p(l,K_{i,j},X_{i,j})}\right), \quad h,l \in \mathcal{H}, \tag{12}$$

where $Z_{i,j}(h,l,k)$ is the log-likelihood ratio defined in Eq. (3). Intuitively, the fact that $S_{i,t}(h,l) > 0$ implies that given the inspections and their outcomes up till time t, the label of job i is *more likely* to be h than l, and such likelihood intensifies as the value of $S_{i,t}(h,l)$ increases. As we will see in a moment, the set $\{S_{i,t}(h,l)\}_{h,l\in\mathcal{H}}$ serves as a summary statistic that is sufficient for producing classifications for job i's label.

Point (b) concerns the quantity of information needed to make an *accurate* classification. In light of the preceding discussion, we could equally ask: at the time when a classification has to be made about job i's label, what conditions should $\{S_{i,t}(h,l)\}_{h,l\in\mathcal{H}}$ satisfy in order for the classification error to be small? The next lemma provides such a sufficient condition. Define $\widehat{H}_{i,t}$ as the *maximum-likelihood (ML) estimator* for the true label of job i, H_i, given the inspections performed on job i up till time t, i.e.,

$$\widehat{H}_{i,t} \in \{h \in \mathcal{H} : S_{i,t}(h,l) \geq 0,\ \forall l \neq h\} = \arg\max_{h\in\mathcal{H}} \prod_{j=1}^{N_{i,t}} p(h,K_{i,j},X_{i,j}), \tag{13}$$

with ties broken arbitrarily[6]. We will denote by $S_i^F(h,l)$ and $\widehat{H}_i^F$ the value of $S_{i,t}(h,l)$ and $\widehat{H}_{i,t}$ at the time when job i departs from the system, respectively. We have the following lemma, which is a special case of a more general result, Lemma A.1, in Appendix A.4.

Lemma 4.1 *Fix $i \in \mathbb{N}$ and $x > 0$. Denote by $\mathcal{G}_x$ the event:*

$$\mathcal{G}_x = \{\exists h' \in \mathcal{H},\ s.t.\ S_i^F(h',l) \geq x, \quad \forall l \in \mathcal{H}, l \neq h'\}. \tag{14}$$

We have that

$$\mathbb{P}(\widehat{H}_i^F \neq h\,,\, \mathcal{G}_x \,\big|\, H_i = h) \leq c_{\mathcal{H}} \exp(-x), \quad \forall h \in \mathcal{H}. \tag{15}$$

Lemma 4.1 shows that if, for a large value of x, the event $\mathcal{G}_x$ occurs with high probability under any job label, then the resulting probability of mis-classification must be small. This achievability result is complemented by the following converse, which states that in order for any inspection policy to be δ-accurate, the expected value of the cumulative log-likelihood ratio, $\mathbb{E}(S_i(h,l))$, must satisfy a lower bound that is approximately $\ln(1/\delta)$. The proof of the lemma utilizes a coupling argument similar to that in Wald (1945) for establishing a lower bound on sample complexity in sequential hypothesis testing, and is given in Appendix A.1.

Lemma 4.2 *Fix $\delta \in (0,1/2)$. If an inspection policy is δ-accurate, then for all $i \in \mathbb{N}$ and $h \in \mathcal{H}$,*

$$\mathbb{E}\left(S_i^F(h,l) \,\big|\, H_i = h\right) \geq (1-\delta)\ln\frac{1-\delta}{\delta} - e^{-1}, \quad \forall l \in \mathcal{H}\backslash\{h\}. \tag{16}$$

The preceding lemmas combined hence give us a more complete picture of the *information requirement* for accurate classification: it suffices that by the time a job i departs from the system, there exists one label, h, whose cumulatively log-likelihood ratio when compared against any other alternative label, l, is sufficiently large, i.e., $S_{i,t}(h,l)$ is at least $\ln(c_{\mathcal{H}}/\delta)$ for all $l \neq h$ (Lemma 4.1), and this is essentially necessary (Lemma 4.2).

[6] Note that the equality in the above equation follows from the definition: the most likely label is also the one that is no less likely than any other labels.

More importantly, the above discussion reveals a natural link from *information need* to *service requirement*. While in a traditional processing system a job may come with a certain *size*, we can think of a job in the information processing system as having a *vector-valued* service requirement: for a job i with true label h, if we interpret the quantity $S_{i,t}(h,l)$ as the amount of "work" already performed along the lth coordinate, then the job's service requirement would be, roughly speaking, that the work performed along *all* coordinates, $l \in \mathcal{H}\backslash\{h\}$, should surpass $\ln(c_{\mathcal{H}}/\delta)$.

This brings us to the last, and arguably most complex, sub-problem, (c): *how* do we satisfy these service requirements in an efficient manner, and simultaneously for multiple jobs? Going back to the definition of $S_{i,t}(h,l)$ in Eq. (12), we see that it can be written as the summation of the $Z_{i,j}(h,l,K_{i,j})$'s. Notably, if job i has true label h, then $\mathbb{E}(Z_{i,1}(h,l,k)) = D(h,l,k)$, where $D(h,l,k)$ is the KL-divergence defined in Eq. (4). In other words, one inspection performed by an expert of type k contributes, in expectation, $D(h,l,k)$ amount of "work" to $S_{i,t}(h,l)$. Viewing our processing task from this angle reveals a resemblance with a certain multi-class multi-server queueing system, where jobs come in different classes (and in this case, labels), and the amount of work a server can contribute to a job's service requirement during a unit time period depends both on the server's type and the type of the job being treated.

Unfortunately, there remains a difficult yet fundamental obstacle that prevents us from directly applying our understanding of multi-class queueing systems to designing inspection policies. The above analogy makes it clear that some expert types may be more informative for a certain job label than others, because the values of $D(h,l,k)$ can vary across h,l, and k. Hence, to best harness the processing power of the experts and satisfy the service requirements across all jobs, the inspections should be arranged in a way that takes into account the types of experts performing the inspections *and* the labels of jobs being inspected, for otherwise an expert could end up wasting her time working on jobs that she has little expertise on. However, this brings us to a circular argument: while efficient inspection beckons a policy to be aware of job labels, we simply do *not* know the job labels, for otherwise there would have not been a need to perform any inspection to begin with!

4.1.1 Overview of the Inspection Policy

Our inspection policy will make use of a three-stage architecture to circumvent the above-mentioned "circular" logic, illustrated in Figure 1.

1. In the first stage (Preparation), the policy "boot-straps" each incoming job, by inspecting it using randomly chosen experts with the goal of generating a *coarse estimate* of its true label.
2. In the second stage (Adaptive), the policy performs the majority of the inspections and in an adaptive manner, with the main goal of verifying whether the coarse estimates are correct. Most of the jobs with a correct coarse estimate will be able to obtain an accurate classification by the end of the Adaptive stage and depart from the system.
3. The third stage (Residual) treats those jobs whose coarse estimates were erroneous to ensure that they, too, will receive an accurate classification.

We will show that (1) the coarse estimates in the Preparation stage are sufficiently accurate so that little resource is wasted in the Adaptive stage, and (2) the processing resources required in the Preparation and Residual stages amount to only a small fraction of the total resources. Together, they lead to the resource efficiency of our inspection policy.

We can also interpret the high-level structure of this three-stage architecture through a *learning* versus *verification* dichotomy: all jobs are first inspected by some "generalists" (i.e., random experts) to *learn* a coarse label estimate. The system then enlists the "specialists" to *verify* the validity of these estimates to a high accuracy. If a coarse estimate is deemed incorrect by the "specialists", the job is then sent *back* to the "generalists" to perform learning thoroughly to reach an accurate estimate, albeit in a less efficient manner.

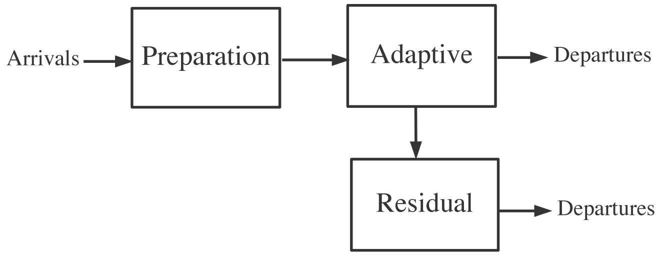

Figure 1: Overall architecture associated with the proposed inspection policy.

4.2 Proof Outline

We now provide a brief outline of the main steps of the proof. Expanding upon the informal discussion in the previous subsection, we formally describe in Section 5 a prior-oblivious inspection policy that will be used to achieve the scaling in Theorem 3.1. In Section 6, we build on Lemma 4.2 and establish a lower bound on the number of experts that *any* δ-accurate policy must satisfy, which is expressed in terms of a solution to a certain linear optimization problem. This lower bound will serve as our performance benchmark of what an "optimal" inspection policy could achieve in terms of minimum system size. In Sections 7 and 8, we establish a sufficient condition for the number of experts under which the proposed policy would stabilize all three stages. In particular, Section 7 contains the most technical portion of our proof, which relies on a fluid model to analyze the joint dynamics of the Preparation and Adaptive stages. We complete the proof of Theorem 3.1 in Section 9, in two steps. We first show that the proposed policy is δ-accurate, which, in light of Lemma 4.1, follows by construction in a straightforward manner. We then compare the sufficient condition on the number of experts, established in Sections 7 and 8, to the lower bound in Section 6, and demonstrate that the ratio between the two converges to 1 uniformly over all prior distributions for the job labels. This completes the proof of Theorem 3.1.

5 Design of the Inspection Policy

We present in this section the inspection policy that we will use to prove Theorem 3.1. The main job-flow of the policy consists of three stages: Preparation, Adaptive, and Residual, as is illustrated in Figure 1. We begin by explaining some basic actions of the experts.

5.1 Randomized Expert Visits

We first introduce some notation. Define

$$r_k = \mu_k \rho_k, \quad k \in \mathcal{K}, \tag{17}$$

and from the assumption in Eq. (2), we have that $\sum_{k\in\mathcal{K}} r_k = \overline{\mu} = 1.$. Because the lengths of inspections are exponentially distributed, r_k is the probability that the next available expert is of type k assuming that all experts are in BUSY in the present moment, and mr_k is the average number of inspections that the pool of type-k experts can complete in unit time. Let $d(h,l)$ be the average KL-divergence for when a job is inspected by an expert whose type is randomly drawn according to the distribution $\{r_k\}_{k\in\mathcal{K}}$:

$$d(h,l) = \sum_{k\in\mathcal{K}} D(h,l,k) r_k. \tag{18}$$

Denote by d_{a} the minimum value among the $d(h,l)$, $d_{\mathrm{a}} = \min_{h,l\in\mathcal{H},h\neq l} d(h,l)$, and by $\overline{z}$ the maximum log-likelihood ratio[7]

$$\overline{z} = \max_{h,l\in\mathcal{H},\,k\in\mathcal{K}} \max_{x:p(l,k,x)\neq 0} \left| \ln \frac{p(h,k,x)}{p(l,k,x)} \right|. \tag{19}$$

Finally, define the constant

$$\zeta_0 = \frac{8\overline{z}^2 + 2d_{\mathrm{a}}}{d_{\mathrm{a}}^2}. \tag{20}$$

Expert visits. We say that an expert of type k *goes on a vacation* to mean that she starts processing a "dummy" job and remains in state BUSY for a period of time that is exponentially distributed with mean $1/\mu_k$. Suppose that an expert completes inspecting a previous job at time t, then she will choose to **visit** one of the three stages, which means that the expert will either initiate an inspection for a job in that stage, or go on a vacation, depending on the inspection rules which will be specified in the next subsection. The choice of which stage to visit will be made by a simple randomized rule, independent of the rest of the system: the expert visits the Preparation, Adaptive, and Residual stage, with probability q^P, q^A, and q^R, respectively, where

$$q^P = \frac{\zeta_0 \ln\ln(1/\delta) + \ln^{-1}(1/\delta)}{m}, \qquad q^R = \frac{3c_{\mathcal{H}}\zeta_0(1+\ln(4c_{\mathcal{H}})\ln^{-1}(1/\delta)) + 1}{m}, \tag{21}$$

$$q^A = 1 - q^P - q^R, \tag{22}$$

and we assume that the policy will be applied under a range of parameters where all expressions above lie in the interval $(0,1)$.

5.2 Multi-stage Inspection Policy

We now describe our inspection policy in detail, where the exposition for each stage is broken down into three parts: (1) Workload: how many inspections need be completed on a job in each stage; (2) Departure rules: where the job goes next; (3) Expert actions: how experts perform inspections.

5.2.1 Preparation Stage

Every job that arrives to the system will first be processed in the Preparation stage. The objective is to produce a coarse estimate of the job's label using only a *small* number of inspections, performed by random experts whose types are drawn according to $\{r_k\}_{k\in\mathcal{K}}$. The randomization in the expert types ensures that information is gained about the job's true label at a non-zero rate. The coarse label estimate will then be used to "bootstrap" processing in the Adaptive stage to further enhance the classification accuracy.

Workload. Every job will receive n^P inspections in the Preparation stage, where

$$n^P = \zeta_0 \ln\ln(1/\delta). \tag{23}$$

Departure rules. When a job has received the *outcomes* from all n^P inspections, it departs from the Preparation stage and enters the Adaptive stage.

Expert actions. Denote by $\mathbf{W}_0(t)$ the total number of *uninitiated* inspections in the Preparation stage at time t. An expert who visits the Preparation stage at time t will attempt to initiate an inspection for a job in the Preparation stage in a first-come-first-serve fashion. If $\mathbf{W}_0(t) = 0$, then the expert goes on a vacation.

[7] The set $\mathcal{X}$ being finite ensures that $\overline{z} < \infty$.

5.2.2 Adaptive Stage

The Adaptive stage is the "power-house" of the system that performs the majority of all inspections. Its defining feature is that the number of inspections that a job receives from each expert type will be decided *adaptively* depending both on the coarse label estimate from the Preparation stage, and on the existing aggregate workload in the Adaptive stage. The main objective of this stage is to *verify* the correctness of the coarse label estimates: *most* jobs with a correct coarse estimate depart from the system after the Adaptive stage, while those with incorrect coarse estimates are likely to be sent to the Residual stage for further processing.

Workload. The workload generation in this stage is more complex than that of the Preparation stage. Upon arriving to the Adaptive stage, a job, i, is assigned a *workload vector* $\{\Lambda_{i,k}\}_{k\in\mathcal{K}}$, where $\Lambda_{i,k}$ is the number of inspections to be performed by experts of type k on job i during its stay in the Adaptive stage. We will denote by $\overline{\Lambda}_{i,k}(t)$ the remaining number of inspections by experts of type k at time t, defined by the difference between $\Lambda_{i,k}$ and the number of inspections already *initiated* by experts of type k for job i by time t in the Adaptive stage. Denote by $Q^A(t)$ the set of jobs in the Adaptive stage at time t. We define the *workload at expert pool* k as

$$\mathbf{W}_k(t) = \sum_{i\in Q^A(t)} \overline{\Lambda}_{i,k}(t), \quad t \in \mathbb{R}_+. \tag{24}$$

We will refer to $\mathbf{W}(\cdot) = \{\mathbf{W}_k(\cdot)\}_{k\in\mathcal{K}}$ as the workload process.

We now explain how the workload vectors, $\{\Lambda_{i,k}\}_{k\in\mathcal{K}}$, are generated. Denote by $\widehat{H}_i^P$ the maximum likelihood estimator of job i, $\widehat{H}_{i,t}$ (Eq. (13)), at the time it exits the Preparation stage, and suppose that $\widehat{H}_i^P = h$. Let $\{n_{h,k}\}_{k\in\mathcal{K}}$ be an optimal solution to the following linear optimization problem:

$$\text{minimize} \quad \sum_{k\in\mathcal{K}} n_k \mathbf{W}_k(t), \tag{25}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}} D(h,l,k) n_k \geq \ln(2c_{\mathcal{H}}/\delta) + g_\delta, \quad \forall l \in \mathcal{H}\backslash\{h\}, \tag{26}$$

$$n_k \geq 0, \quad \forall k \in \mathcal{K}, \tag{27}$$

$$\sum_{k\in\mathcal{K}} n_k \leq v_\delta, \tag{28}$$

with ties broken arbitrarily, where g_δ and v_δ are two auxiliary constants that do not depend on h:

$$g_\delta = 3\overline{z}\sqrt{c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\ln\ln(1/\delta)} \tag{29}$$

$$v_\delta = 2c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\left[1 + (\ln(2c_{\mathcal{H}}) + g_\delta)\ln^{-1}(1/\delta)\right]. \tag{30}$$

We will denote by $\mathcal{N}_h$ the set of all vectors $\{n_k\}_{k\in\mathcal{K}}$ which satisfy the constraints in Eqs. (26) through (28). One can verify that the above optimization problem always admits a feasible solution. Finally, the workload vector for job i will be obtained by rounding down $\{n_{h,k}\}_{k\in\mathcal{K}}$:

$$\Lambda_{i,k} = \lfloor n_{h,k} \rfloor, \quad \forall k \in \mathcal{K}. \tag{31}$$

Interpretation of workload: The workload vector captures the combinations of inspections job i should receive *assuming* that its true label is indeed $\widehat{H}_i^P$, in which case Eq. (26) ensures that the cumulative log-likelihoood ratios are sufficiently large to make an accurate classification. As mentioned earlier, the optimization in Eq. (25) is reminiscent of the family of max-weight scheduling policies (cf. Tassiulas and Ephremides (1992)): our policy aims to create workload in order to minimize an inner-product between the new workload and the existing inspections. The proof in subsequent sections will demonstrate that this adaptive procedure allows the inspection policy to work well with any prior distribution, hence making the policy prior-oblivious.

Departure rules. A job i departs from the Adaptive stage as soon as it has received the outcomes of all $\Lambda_{i,k}$ inspections from experts of type k, for all $k \in \mathcal{K}$. Suppose that the departure occurs at time t. The policy then executes the following decision:

1. If there exists $h \in \mathcal{H}$, such that

$$S_{i,t}(h,l) \geq \ln(2c_{\mathcal{H}}/\delta), \quad \forall l \in \mathcal{H}, l \neq h, \tag{32}$$

   then job i departs from the system, and a classification is produced by setting $\widehat{H}_i = \widehat{H}_{i,t} = h$.

2. Otherwise, job i enters the Residual stage.

Expert actions. Suppose that an expert of type k visits the Adaptive stage at time $t \in \mathbb{R}_+$. If the workload for the kth expert pool, $\mathbf{W}_k(t)$, is non-zero (Eq. (24)), then the expert initiates an inspection for a job associated with one unit of work in $\mathbf{W}_k(t)$, in a first-come-first-serve fashion. If $\mathbf{W}_k(t) = 0$, then the expert goes on a vacation.

5.2.3 Residual Stage

The Residual stage acts as a "clearing house" that treats those jobs whose inspections in the Adaptive stage failed to produce an accurate classification. Similar to the Preparation stage, jobs are inspected by random experts, but they receive significantly more inspections in the Residual stage in order to produce a highly accurate label classification.

Workload. The moment a job enters the Residual stage, all of its previous inspections and outcomes are *discarded*. Similar to the Preparation stage, each job will receive a fixed number of inspections, n^R, where

$$n^R = \zeta_0 \ln(4c_{\mathcal{H}}/\delta). \tag{33}$$

Departure rules. A job i in the Residual stage departs from the system as soon as it has received the results from all n^R inspections, and a classification of job i's type is produced by setting $\widehat{H}_i = \widehat{H}_{i,t}$. Note that the classification is made solely based on the inspections in the Residual stage, as we have discarded all inspections from the earlier stages.

Expert actions. An expert who visits the Residual stage will attempt to initiate an inspection for a job in the Residual stage in a first-come-first-serve fashion. If there is no job currently in the Residual stage, or if all jobs in the Residual stage have all of their n^R inspections already initiated, then the expert goes on a vacation. This concludes the description of our inspection policy.

6 Lower Bound on Optimal System Size

We establish in this section a fundamental lower bound on the minimum number of experts required in order for the system to be stable that holds for *any* δ-accurate policy. The following Fundamental Linear Program is central to this lower bound as well as our subsequent analysis.

Definition 6.1 *The* **Fundamental Linear Program***, denoted by* FLP*, is defined as follows.*

$$\begin{aligned}
\textit{minimize} \quad & m && (34)\\
\textit{s.t.} \quad & \sum_{h\in\mathcal{H}} n_{h,k}\pi_h \leq r_k m, \quad k \in \mathcal{K}, && (35)\\
& \sum_{k\in\mathcal{K}} n_{h,k} D(h,l,k) \geq \ln(1/\delta), \quad \forall h,l \in \mathcal{H}, h \neq l, && (36)\\
& n_{h,k} \geq 0, \quad \forall h \in \mathcal{H}, k \in \mathcal{K}, && (37)
\end{aligned}$$

where $D(h,l,k)$ and r_k are defined in Eqs. (4) and (17), respectively.

We provide some intuition to motivate the above definition. Recall from Lemma 4.2 that in order for any policy to be δ-accurate, for a job i with label h, the expected value of the cumulative log-likelihood ratio $S_{i,t}(h,l)$ should be at least $\ln(1/\delta)$ by the time job i departs from the system. Furthermore, an inspection by an expert of type k increases the value of $S_{i,t}(h,l)$ by $D(h,l,k)$ in expectation. If we interpret the variables, $n_{h,k}$, as the number of inspections a job with true label h should receive from an expert of type k, then Eq. (35) in FLP corresponds to the resource constraint of having $m\rho_k$ type-k experts, each of whom can perform μ_k inspections per unit time, and Eq. (36) to the above-mentioned constraint on $\mathbb{E}(S_{i,j}(h,l))$ imposed by Lemma 4.2. Therefore, FLP captures the problem of finding minimal system size faced by a decision maker who already *knows* the true labels of the jobs and is only interested in "verifying" them in order to satisfy the condition of Lemma 4.2, and we would expect the optimal value of FLP to be a lower bound for what is achievable in our problem, where the job labels are unknown.

The following proposition is the main result of this subsection, which states that the minimum system size under any δ-accurate policy is essentially no smaller than the solution to FLP, as $\delta \to 0$. The proof builds upon the lower bound in Lemma 4.2, and is given in Appendix A.2.

Proposition 6.2 *Fix π. Denote by m_F^* the optimal value of* FLP. *There exists $\delta_F \in (0,1)$, such that for all $\delta \in (0, \delta_F)$,*

$$m^*(\delta,\pi) \geq b_\delta \cdot m_F^*, \tag{38}$$

where $b_\delta = (1-\delta)\left[1 - \left(\ln \frac{1}{1-\delta} + e^{-1}\right)\ln^{-1}(1/\delta)\right]$. In particular, $b_\delta \uparrow 1$ as $\delta \downarrow 0$.

The next lemma states some useful properties of FLP. The proof is given in Appendix A.3.

Lemma 6.3 *There exists an optimal solution of* FLP, *$(m_F^*, \{n_{h,k}^*\})$, such that the following holds:*

$$\overline{d}^{-1} \leq \frac{1}{\ln(1/\delta)} \sum_{k\in\mathcal{K}} n_{h,k}^* \leq c_{\mathcal{K}} \underline{d}^{-1}, \quad \forall h \in \mathcal{H}, \tag{39}$$

$$m_F^* \geq \overline{d}^{-1} \ln(1/\delta). \tag{40}$$

7 Stability of Preparation and Adaptive Stages

We establish in this section a sufficient condition on the number of experts, m, in order for the Preparation and Adaptive stages to be jointly stable under the proposed inspection policy. We show that the joint dynamics in the two stages are captured by a countable-state Markov process that is positive recurrent. The main result of this section is the following theorem.

Theorem 7.1 *Define $\underline{r} = \min_{k\in\mathcal{K}} r_k = \min_{k\in\mathcal{K}} \rho_k\mu_k$. The Preparation and Adaptive stages are stable whenever $mq^P > n^P = \zeta_0 \ln\ln(1/\delta)$, and*

$$mq^A > \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \frac{2c_{\mathcal{H}}^2 c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1 + \ln^{-1}(1/\delta))m_F^*, \tag{41}$$

where m_F^* is the optimal value of the Fundamental Linear Program in Definition 6.1.

Proof Overview for Theorem 7.1. The remainder of this section is devoted to the proof of Theorem 7.1, and, unless stated otherwise, we will use the word "system" to refer to the Preparation and Adaptive stages only. We begin by giving an overview of the proof and highlighting some of the main technical challenges that motivate our approach.

Let us first recall some high-level features of the system dynamics. The expert actions are fairly simple in both stages by simply trying to initiate an inspection in a non-adaptive manner. Creating inspection workloads is also straightforward for the Preparation stage where each job has a fixed number of inspections. The main complexity therefore lies in how the vector-valued workloads are created in the Adaptive stage, which depends both on the job's type estimate from the Preparation stage, and the aggregate workloads in the Adaptive stage. Given the disparity of complexity, a natural approach would be to treat the two stages separately: the Preparation stage admits simpler dynamics and is easy to analyze while the Adaptive could be tackled using the Foster-Lyapunov criterion. Unfortunately, this approach falls short because the processing in the Preparation stage destroys the memoryless property of the initial arrival process, rendering its output process non-Markovian. Therefore, the Adaptive stage cannot be treated as an isolated Markov process without taking into account the state of the Preparation stage as well.

To overcome this problem, we will model the dynamics in both stages *jointly*, and formulate a set of *fluid solutions*, expressed as solutions to a system of ordinary differential equations (ODE), to capture the essential dynamics of this joint process. Specifically, following a general program developed by Rybko and Stolyar (1992) and Dai (1995), we will show that, under proper scaling, the process of system workloads converges almost surely to a set of fluid solutions. We then show that if the number of experts satisfies the conditions stated in Theorem 7.1, then the fluid solutions exhibit a certain global contraction property with respect to a one-homogeneous Lyapunov function. These two properties will then be used to show that the original workload process is positive recurrent, which in turn implies the stability of the system. The proof will be carried out in the following main steps:

1. We define in Section 7.2 the Markov process that captures the system dynamics, as well as the fluid solutions.
2. We show in Section 7.3 (Proposition 7.4) that the workload process converges to a fluid solution almost surely over an appropriately defined probability space.
3. We establish the global contraction property of the fluid solutions in Section 7.4 (Proposition 7.11) by showing a multiplicative decay in a quadratic Lyapunov function along the trajectory of *any* fluid solution.
4. Finally, we complete the proof in Section 7.5 by combining Propositions 7.4 (convergence) and 7.11 (contraction) with a variant of the Foster-Lyapunov criterion (Theorem 8.13 of Robert (2003)).

There are two main technical challenges which we develop novel tools to overcome: (1) the proof requires characterizing the limit points of the solutions to the linear optimization subroutine in Eq. (25) under fluid scaling, which is difficult because the parameters in the objective function, $\{\mathbf{W}_k(\cdot)\}_{k\in\mathcal{K}}$, are themselves stochastic variables. We will employ a careful analysis of the (semi-)continuity properties of the optimization problem to study these limit points; (2) the jobs' transitions from the Preparation stage to the Adaptive stage are complex because a job can depart from the Preparation stage only after all of its inspection have completed (for otherwise the inspection outcomes would not have been available). This in turn causes the order of departures from a stage to deviate from that of the arrivals, making standard techniques ineffective in showing the convergence of the system's stochastic trajectory to a fluid limit. We will prove convergence by developing explicit bounds on the degree of "shuffling," which allows us to conclude that the deviation is not too substantial to invalidate convergence.

7.0.1 Additional Notation

For a vector $x = (x_1, \ldots, x_n)$, we will denote by $\|x\|_2$ the l_2 norm of x: $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$. We will denote by $\|\cdot\|_T$ the maximal norm of a function over the interval $[0, T]$: for a function $f : \mathbb{R}_+ \to \mathbb{R}^K$, $\|f(t)\|_T = \sup_{t\in[0,T]} \max_{1\le k\le K} |f_k(t)|$. For $x, y \in \mathbb{R}$, we will use $x \vee y$ and $x \wedge y$ to denote $\max\{x, y\}$ and $\min\{x, y\}$, respectively. For two vectors of the same dimension, $\mathbf{x}$ and $\mathbf{y}$, we write $\mathbf{x} \le \mathbf{y}$ if all coordinates of $\mathbf{x}$ are dominated by those of $\mathbf{y}$. Similarly, for a set of vectors, $\mathcal{Y}$, we write $\mathbf{x} \le \mathcal{Y}$ if $\mathbf{x} \le \mathbf{y}$ for all $\mathbf{y} \in \mathcal{Y}$. The addition of two sets $\mathcal{X} + \mathcal{Y}$ is defined to be the set $\{z : z = x + y, x \in \mathcal{X}, y \in \mathcal{Y}\}$.

7.1 Classification Accuracy of Preparation Stage

We begin in this subsection with a result that bounds the classification error of the coarse label estimate from the Preparation stage. Denote by $\widehat{H}_i^P$ the maximum likelihood estimator, $\widehat{H}_{i,t}$, when job i exits the Preparation stage. Recall from the construction of our policy that each job will be inspected for the same, deterministic number of times in the Preparation stage, and it is not difficult to verify that $\{\widehat{H}_i^P\}_{i\in\mathbb{N}}$ are i.i.d. We will denote by π^P the distribution of the estimator $\widehat{H}_1^P$,

$$\pi_h^P \triangleq \mathbb{P}\left(\widehat{H}_1^P = h\right), \quad h \in \mathcal{H}. \tag{42}$$

and by ϵ^P as the error probability $\epsilon^P \triangleq \max_{h\in\mathcal{H}} \mathbb{P}(\widehat{H}_1^P \neq h \,|\, H_1 = h)$. The following proposition provides an upper bound on ϵ^P, which in turn upper-bounds the distance from π^P to the original prior distribution, π. The proof is given in Appendix A.5, which relies on a generalization of Lemma 4.1 in combination with fact that $S_{i,t}(h,l)$ can be viewed as a martingale under proper conditioning.

Proposition 7.2 *We have that* $\epsilon^P \leq 2c_{\mathcal{H}} \ln^{-1}(1/\delta)$, *and*

$$\pi_h^P \leq \pi_h + \epsilon^P \sum_{h'\neq h} \pi_{h'} \leq \pi_h + 2c_{\mathcal{H}} \ln^{-1}(1/\delta), \quad \forall h \in \mathcal{H}. \tag{43}$$

7.2 State Representation and Fluid Solutions

We now describe a Markovian representation of the Preparation and Adaptive stages as well as the notion of fluid solutions. We will index jobs in the two stages according to the order in which they arrive. Denote by $I(t)$ the total number of jobs in the system at time t, and let $\mathcal{I}(t) = \{1, 2, \ldots, I(t)\}$. For instance, job 1 corresponds to the oldest job in system at time t, and job $I(t)$ corresponds to the youngest. The indices will be updated accordingly in the event of the departure of job i, where all jobs with an index greater than i will have their index reduced by 1. For each $i \in \mathcal{I}(t)$, we define a *job state*, $Y_i(t)$, which consists of the following variables:

1. $Y_i^S(t) \in \{1, 2\}$ represents the stage the job is currently in, with 1 and 2 corresponds to job i being in the Preparation and Adaptive stage, respectively.
2. $\{(X_{i,s}, E_{i,s})\}_{s=1}^{N_{i,t}}$ contains all the past inspection responses of job i, along with the corresponding expert types, where $N_{i,t}$ the number of inspections received by job i by time t.
3. $Y_i^E(t) \subset \mathcal{E}$ is the set of experts who are in the process of inspecting job i at time t.
4. $N_i^P(t)$ is the number of remaining inspections that job i has left in the Preparation stage.
5. $\overline{\Lambda}_{i,k}(t)$ is job i's remaining number of inspections to be completed by experts of type k in the Adaptive stage.

The above variables completely specify the state of the Preparation and Adaptive stages at time t, and it is not difficult to verify that $\{\mathcal{I}(t), \{Y_i(t)\}_{i\in\mathcal{I}(t)}\}_{t\in\mathbb{R}_+}$ is an irreducible countable-state Markov process. We now formally define the *workload process* using the above state representation. Denote by $\mathbf{W}_0(t)$ the total number of remaining uninitiated inspections in the Preparation stage

$$\mathbf{W}_0(t) = \sum_{i\in\mathcal{I}(t), Y_i^S(t)=1} N_i^P(t). \tag{44}$$

Similarly, the *workload in the kth expert pool* in the Adaptive stage, is defined by:

$$\mathbf{W}_k(t) = \sum_{i\in\mathcal{I}(t), Y_i^S(t)=2} \overline{\Lambda}_{i,k}(t). \tag{45}$$

The following notion of fluid solutions will serve as an approximation to the workload process.

Definition 7.3 (Fluid Solutions) *The functions* $\mathbf{w}_k : \mathbb{R}_+ \to \mathbb{R}_+, k \in \{0\} \cup \mathcal{K}$ *are called a* fluid solution *if there exist Lipschitz-continuous functions* $\mathbf{a}_k, \mathbf{d}_k : \mathbb{R}_+ \to \mathbb{R}_+$, $k \in \{0\} \cup \mathcal{K}$, *with* $\mathbf{a}_k(0) = \mathbf{d}_k(0) = 0$ *with Lipschitz constant* $c_{\mathcal{L}} > 0$ *such that*

$$\mathbf{w}_k(t) = \mathbf{w}_k(0) + \mathbf{a}_k(t) - \mathbf{d}_k(t), \tag{46}$$

where, for almost all $t \in \mathbb{R}_+$,

$$\dot{\mathbf{a}}_0(t) = n^p, \qquad \dot{\mathbf{d}}_0(t) = \begin{cases} mq^P, & \text{if } \mathbf{w}_0(t) > 0, \\ n^P, & \text{if } \mathbf{w}_0(t) = 0, \end{cases}$$

$$\{\dot{\mathbf{a}}_k(t)\}_{k \in \mathcal{K}} \leq (1 + \ln^{-1}(1/\delta)) \sum_{h \in \mathcal{H}} \pi_h^P \mathcal{N}_h^*(\mathbf{w}(t)),$$

$$\dot{\mathbf{d}}_k(t) = \begin{cases} r_k mq^A, & \text{if } \mathbf{w}_k(t) > 0, \\ \dot{\mathbf{a}}_k(t), & \text{if } \mathbf{w}_k(t) = 0, \end{cases} \qquad \forall k \in 1, \ldots, c_{\mathcal{K}}, \tag{47}$$

where q^P, q^A *and* n^P *were defined in Section 5, and* π_h^P *in Eq. (42).* $\mathcal{N}_h^*(\mathbf{w}(t))$ *is the set of optimal solutions for the optimization problem*

$$\min_{\{n_k\} \in \mathcal{N}_h} \sum_{k \in \mathcal{K}} n_k \mathbf{w}_k(t), \tag{48}$$

and the set $\mathcal{N}_h$ *was defined in Eqs. (26) through (28). Fix* $\mathbf{w}^0 \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$. *We denote by* $\mathcal{W}(\mathbf{w}^0)$ *the set of all fluid solutions with the initial condition* $\mathbf{w}(0) = \mathbf{w}^0$.

7.3 Convergence of Stochastic Sample Paths to the Fluid Solutions

We show in this section that the workload process converges to a fluid solution under proper scaling. We will consider a sequence of systems, indexed by $s \in \mathbb{N}$, which have different initial conditions for the workload process at $t = 0$ but are otherwise identical. We will denote by $\mathbf{W}^s(\cdot)$ the workload process $\mathbf{W}(\cdot)$ in the sth system. For $n \in \mathbb{N}$, and a process $\{X(t)\}_{t \in \mathbb{R}_+}$, we will use $\{X(n,t)\}_{t \in \mathbb{R}_+}$ to denote the *normalized process*:

$$X(n,t) = \frac{1}{n} X(nt), \quad t \in \mathbb{R}_+. \tag{49}$$

The following proposition is the main result of this subsection.

Proposition 7.4 *Fix* $\mathbf{w}^0 \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$. *Consider a sequence of initial conditions* $\{\mathbf{w}^{(s)}\}_{s \in \mathbb{N}}$, *such that for a sequence of positive numbers* $\{z_s\}_{s \in \mathbb{N}}$ *with* $\lim_{s \to \infty} z_s = \infty$, *we have that* $\lim_{s \to \infty} z_s^{-1} \mathbf{w}^{(s)} = \mathbf{w}^0$.

Suppose $\mathbf{W}^s(0) = \mathbf{w}^{(s)}$, *for all* $s \in \mathbb{N}$. *Then, for all* $T > 0$, *the following convergence takes place:*

$$\lim_{s \to \infty} \inf_{\mathbf{w} \in \mathcal{W}(\mathbf{w}^0)} \|\mathbf{W}^s(z_s, t) - \mathbf{w}(t)\|_T = 0, \quad \textit{almost surely.} \tag{50}$$

The remainder of this sub-section is devoted to the proof of Proposition 7.4. Because the system dynamics is quite complex, we will prove convergence using a sample-path-based approach that helps us isolate the probabilistic aspect of the dynamics from its deterministic counterpart (Similar sample-path-based approaches have also been used in Bramson (1998) and Tsitsiklis and Xu (2012)). In particular, we first identify a large subset of the sample space, called the *regular set*, which contains the sample paths that exhibit certain typical behaviors. We then show that convergence to the fluid solutions occurs over *every* sample path in the regular set.

7.3.1 Regular Set

We now construct the regular set. We will define all random quantities of the system on the same probability space, $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 7.5 *Fix $T > 0$. We define the following elements of $\mathcal{F}$.*

1. *$\mathcal{C}_0$: Denote by $\Xi_0(t)$ the number of jobs that have arrived to the Preparation stage in the interval $[0, t]$. Define $\mathcal{C}_0$ as the event where*

$$\lim_{z\to\infty} \|\Xi_0(z,t) - t\|_T = 0, \tag{51}$$

2. *$\mathcal{C}_\mathcal{E}$: Denote by $R_{e,P}(t)$ and $R_{e,A}(t)$ the total number of times an expert $e \in \mathcal{E}$ visits the Preparation and Adaptive stage, respectively, during the interval $[0, t]$. Let k_e be the type of expert e. Define $\mathcal{C}_\mathcal{E}$ as the event where*

$$\lim_{z\to\infty} \max_{e\in\mathcal{E}} \left\|R_{e,P}(z,t) - q^P \mu_{k_e} t\right\|_T = 0, \quad \text{and} \quad \lim_{z\to\infty} \max_{e\in\mathcal{E}} \left\|R_{e,A}(z,t) - q^A \mu_{k_e} t\right\|_T = 0. \tag{52}$$

3. *$\mathcal{C}_H$: Denote by $\widehat{H}_i^P$ the ML estimator of the type of the ith job upon leaving the Preparation stage. Let $g_h(t) = \sum_{i=1}^{\lfloor t \rfloor} \mathbb{I}(\widehat{H}_i^P = h)$. Define $\mathcal{C}_H$ as the event where*

$$\lim_{z\to\infty} \max_{h\in\mathcal{H}} \left\|g_h(z,t) - \pi_h^P t\right\|_T = 0. \tag{53}$$

Define the *regular set*, $\mathcal{C}$, as the intersection of all three events in Definition 7.5: $\mathcal{C} = \mathcal{C}_0 \cap \mathcal{C}_\mathcal{E} \cap \mathcal{C}_H$. We have the following useful lemma. The proof is a direct consequence of the (functional) law of large numbers applied to each of the three events, and is omitted.

Lemma 7.6 *Fix $T > 0$. We have that $\mathbb{P}(\mathcal{C}_0) = \mathbb{P}(\mathcal{C}_\mathcal{E}) = \mathbb{P}(\mathcal{C}_H) = 1$, and, consequently, $\mathbb{P}(\mathcal{C}) = 1$.*

7.3.2 Proof of Proposition 7.4

We return to the proof Proposition 7.4, which will be completed in two parts. Fix any sample path $\omega \in \mathcal{C}$. In the first part, we will show that over any finite interval, any sub-sequence of $\{\mathbf{W}^{z_s}(\cdot)\}_{s\in\mathbb{N}}$ admits a further converging sub-sequence that converges coordinate-wise to a Lipschitz-continuous function. We will then show, in the second part, that all such limiting functions are in fact fluid solutions, i.e., for almost all t, their derivatives satisfy the conditions given in Definition 7.3. The next result summarizes the first part of the proof. Fix $k \in \{0, 1, \ldots, c_\mathcal{K}\}$. We will write

$$\mathbf{W}_k^s(t) = \mathbf{W}_k^s(0) + \mathbf{A}_k^s(t) - \mathbf{\Delta}_k^s(t), \quad t > 0, \tag{54}$$

where $\mathbf{A}_k^s(t)$ and $\mathbf{\Delta}_k^s(t)$ denote the total number of inspections associated with the arrivals, and the number of inspections that have been initiated, during $[0, t]$, respectively, for jobs associated with the workload $\mathbf{W}_k^s(\cdot)$. The proof of the following proposition is given in Appendix A.6.

Proposition 7.7 *Fix $T > 0$. Denote by $\mathcal{L}_c$ the set of coordinate-wise c-Lipschitz functions from $[0, T]$ to $\mathbb{R}_+^{3(c_\mathcal{K}+1)}$. Fix the sample, $\omega \in \mathcal{C}$, and an increasing sequence, $\{s_i\}_{i\in\mathbb{N}}$, and let $c = [q^A + (v_\delta + 1)q^P]m + n^P$, where v_δ was defined in Eq. (30). Then, there exists $(\mathbf{w}, \mathbf{a}, \mathbf{d}) \in \mathcal{L}_c$, and an increasing sequence, $\{i_j\}_{j\in\mathbb{N}} \subset \mathbb{N}$, such that*

$$\lim_{j\to\infty} \left\|(\mathbf{W}, \mathbf{A}, \mathbf{\Delta})^{s_{i_j}}(z_{s_{i_j}}, t) - (\mathbf{w}, \mathbf{a}, \mathbf{d})(t)\right\|_T = 0. \tag{55}$$

We will refer to these $\mathbf{w}(\cdot)$'s as limit points of $\{\mathbf{W}^s(z_{s_i}, \cdot)\}_{i\in\mathbb{N}}$.

The next result states that *all* of the limit points in Proposition 7.7 are in fact fluid solutions.

Proposition 7.8 *Let $\mathbf{w}(\cdot)$ be a limit point as defined in Proposition 7.7. Then $\mathbf{w}(\cdot)$ is also a fluid solution, as defined by Definition 7.3.*

Proof. Fix $\omega \in \mathcal{C}$, and a limit point, $(\mathbf{w}, \mathbf{a}, \mathbf{d})$, with the corresponding sequence $\{i_j\}_{j\in\mathbb{N}}$, as defined in Proposition 7.7. To avoid excessive use of subscripts, we will use $\bar{s}_j$ and $\bar{z}_j$ in place of s_{i_j} and $z_{s_{i_j}}$, respectively. Fix $t \in (0,T)$ to be a time where all coordinates of the limit point are differentiable. We begin with the Preparation stage, with $k=0$. As was mentioned in the proof of Proposition 7.7, each new job that arrives to the Preparation stage creates n^P inspections. We have that

$$\lim_{z\to\infty} \left\|\mathbf{A}_0(z,t) - n^P t\right\|_T = \lim_{z\to\infty} n^P \left\|\Xi_0(z,t) - t\right\|_T = 0, \tag{56}$$

where the second equality follows from Eq. (51). This shows that $\dot{\mathbf{a}}_0(t) = n^P$.

For $\mathbf{d}_0(t)$, we consider two cases depending on the value of $\mathbf{w}_0(t)$. First, suppose that $\mathbf{w}_0(t) > 0$. Then there exists $\bar{\epsilon} > 0$, such that for all $\epsilon \in (0, \bar{\epsilon})$, there exists $N_\epsilon > 0$ such that, for all $j \geq N_\epsilon$,

$$\mathbf{W}^{\bar{s}_j}(t'\bar{z}_j) > 0, \quad \forall t' \in [t, t+\epsilon]. \tag{57}$$

Fix j and t' so that Eq. (57) is true. Because the workload is non-zero, any expert who visits the Preparation stage at time $t'\bar{z}_j$ will necessarily lead to a unit increment in the process $\mathbf{\Delta}_0^{\bar{s}_j}(\cdot)$. We thus have that, for all $j \geq N_\epsilon$,

$$\mathbf{\Delta}_0^{\bar{s}_j}((t+\epsilon)\bar{z}_j) - \mathbf{\Delta}_0^{\bar{s}_j}(t\bar{z}_j) = \sum_{e\in\mathcal{E}} R_{e,P}((t+\epsilon)\bar{z}_j) - R_{e,P}(t\bar{z}_j), \tag{58}$$

which, with scaling, implies that

$$\lim_{j\to\infty} (\mathbf{\Delta}_0^{\bar{s}_j}(\bar{z}_j, t+\epsilon) - \mathbf{\Delta}_0^{\bar{s}_j}(\bar{z}_j, t)) = \lim_{j\to\infty} \left(\sum_{e\in\mathcal{E}} R_{e,P}(\bar{z}_j, t+\epsilon) - R_{e,P}(\bar{z}_j, t)\right) = q^P m\epsilon, \tag{59}$$

where the last step follows from Eq. (52), and the fact that $\sum_k \rho_k \mu_k = \overline{\mu} = 1$ (Eq. (2)). Taking the limit as $\epsilon \to 0$, we obtain

$$\dot{\mathbf{d}}_0(t) = \lim_{\epsilon\to 0} \epsilon^{-1} \left[\lim_{j\to\infty} (\mathbf{\Delta}_0^{\bar{s}_j}(\bar{z}_j, t+\epsilon) - \mathbf{\Delta}_0^{\bar{s}_j}(\bar{z}_j, t))\right] = \lim_{\epsilon\to 0} \epsilon^{-1}(q^P m\epsilon) = q^P m, \quad \text{if } \mathbf{w}_0(t) > 0. \tag{60}$$

Now suppose that $\mathbf{w}_0(t) = 0$. In this case, we will exploit the properties that $\mathbf{w}_0(\cdot)$ is differentiable at t, and is *non-negative* over $[0,T]$. Since $\mathbf{w}_0(t) = 0$, the two properties together imply $\dot{\mathbf{w}}_0(t)$ must be zero. We have that $\dot{\mathbf{d}}_0(t) = -(\dot{\mathbf{w}}_0(t) - \dot{\mathbf{a}}_0(t)) = n^P$, if $\mathbf{w}_0(t) = 0$. This proves the case for $k = 0$.

We next consider the case of $k = 1, \ldots, c_{\mathcal{K}}$. For $\dot{\mathbf{d}}_k(t)$, the analysis is identical to the case of $k = 0$, which we shall omit. We now turn to the analysis for $\dot{\mathbf{a}}_k(\cdot)$. In particular, we will show that

$$\{\dot{\mathbf{a}}_k(t)\}_{k\in\mathcal{K}} \leq (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P \mathcal{N}_h^*(\mathbf{w}(t)), \tag{61}$$

where $\mathcal{N}_h^*(\mathbf{w}(t))$ was defined in Eq. (48). The proof of Eq. (61) is more involved than for the other coordinates of the fluid solution, because our inspection policy determines the workload vector of a job entering the Adaptive stage using two sources of information: (1) the coarse estimate of the job's label from the Preparation stage, and (2) the existing workload in the Adaptive stage. Our proof will also proceed in two steps:

(a) We first show that the arrival process of jobs with the same coarse label estimate converges locally to its "mean value" under fluid scaling.

(b) We then show that the average workload vector among jobs with the same coarse estimator converges locally to a point that is dominated by the set $(1 + \ln^{-1}(1/\delta))\mathcal{N}_h^*(\mathbf{w}(t))$.

These two steps together then yield Eq. (61).

For $t \in (0,T)$, denote by $\mathcal{B}^s(t)$ the set of all jobs that arrived to the Adaptive stage during the interval $[0,t]$ in the s-th system, and by $\mathcal{B}^s_h(t)$ the subset of jobs in $\mathcal{B}^s(t)$ whose ML estimators upon exiting the Preparation stage, $\widehat{H}^P_i$, are equal to h. Let $B^s_h(t)$ be the size of the set $\mathcal{B}^s_h(t)$. The following lemma formalizes step (a) above, whose proof is given in Appendix A.7. As was mentioned in the beginning of Section 7, the proof of this lemma involves a careful analysis of the potential shuffling in the order in which jobs depart from the Preparation stage.

Lemma 7.9 *Fix $h \in \mathcal{H}$. For almost all $t \in (0,T)$, we have that*

$$\dot{\mathbf{b}}_h(t) \triangleq \lim_{\epsilon\to 0}\lim_{j\to\infty} \frac{B^{\bar{s}_j}_h(\bar{z}_j, t+\epsilon) - B^{\bar{s}_j}_h(\bar{z}_j,t)}{\epsilon} = \begin{cases} \frac{mq^P}{n^P}\pi^P_h, & \text{if } \mathbf{w}_0(t) > 0,\\ \pi^P_h, & \text{if } \mathbf{w}_0(t) = 0,\end{cases} \tag{62}$$

where $\pi^P_h \triangleq \mathbb{P}\left(\widehat{H}^P_1 = h\right)$, as was defined in Eq. (42).

Step (b) is summarized in the following lemma. It states that, over a small time interval around t, the average workload among the jobs in $\mathcal{B}^{\bar{s}_j}_h(\cdot)$ stays close to a point that is dominated by the set $\mathcal{N}^*_h(\mathbf{w}(t))$. The proof, which utilizes a semi-continuity property of the solution set of the workload-generating optimization problem (Eq. (25)), is given in Appendix A.8.

Lemma 7.10 *Fix $t \in (0,T)$, and $h \in \mathcal{H}$. Let $\overline{\Lambda}_h(t,\epsilon,j)$ be the average workload across all jobs arriving to the Adaptive stage during $[\bar{z}_j t, \bar{z}_j(t+\epsilon))$ whose ML estimator is h, i.e.,*

$$\overline{\Lambda}_h(t,\epsilon,j) = \frac{1}{B^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon)) - B^{\bar{s}_j}_h(\bar{z}_j t)} \sum_{i\in\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon))\setminus\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j t)} \Lambda_i. \tag{63}$$

We have that

$$\limsup_{\epsilon\downarrow 0}\limsup_{j\to\infty} \inf_{y\le\mathcal{N}^*_h(\mathbf{w}(t))} \|\overline{\Lambda}_h(t,\epsilon,j) - y\|_2 = 0. \tag{64}$$

We are now ready to complete the proof of Proposition 7.8. Let $t \in (0,T)$ be a point where all coordinates of $\mathbf{a}(\cdot)$ are differentiable. We have that

$$\begin{aligned}
(\dot{\mathbf{a}}_k(t))_{k=1,\dots,c_\mathcal{K}} &= \lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\frac{\mathbf{A}^{\bar{s}_j}(\bar{z}_j,t+\epsilon) - \mathbf{A}^{\bar{s}_j}(\bar{z}_j,t)}{\epsilon} \overset{(a)}{=} \lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\frac{1}{\epsilon\bar{z}_j}\sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon))\setminus\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j t)}\Lambda_i\\
&= \lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\sum_{h\in\mathcal{H}}\frac{B^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon)) - B^{\bar{s}_j}_h(\bar{z}_j t)}{\epsilon\bar{z}_j}\cdot\frac{\sum_{i\in\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon))\setminus\mathcal{B}^{\bar{s}_j}_h(\bar{z}_j t)}\Lambda_i}{B^{\bar{s}_j}_h(\bar{z}_j(t+\epsilon)) - B^{\bar{s}_j}_h(\bar{z}_j t)}\\
&= \lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\sum_{h\in\mathcal{H}}\frac{B^{\bar{s}_j}_h(\bar{z}_j,t+\epsilon) - B^{\bar{s}_j}_h(\bar{z}_j,t)}{\epsilon}\,\overline{\Lambda}_h(t,\epsilon,j)\\
&\overset{(b)}{=} \lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\sum_{h\in\mathcal{H}}\dot{\mathbf{b}}_h(t)\,\overline{\Lambda}_h(t,\epsilon,j),
\end{aligned} \tag{65}$$

where step (a) follows from the definition of the policy in Eq. (31): $\mathbf{A}^s_k(t) = \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{B}^s_h(t)}\Lambda_{i,k}$, and step (b) from Lemma 7.9. Applying Lemma 7.10 for every $h \in \mathcal{H}$, we have that

$$\limsup_{\epsilon\downarrow 0}\limsup_{j\to\infty}\inf_{y\le\sum_{h\in\mathcal{H}}\dot{\mathbf{b}}_h(t)\,\mathcal{N}^*_h(\mathbf{w}(t))}\left\| y - \sum_{h\in\mathcal{H}}\dot{\mathbf{b}}_h(t)\,\overline{\Lambda}_h(t,\epsilon,j)\right\|_2 = 0. \tag{66}$$

Since $\mathcal{N}_h$, the feasible solutions of the linear program in Eq. (25), is a compact set, the set of optimal solutions, $\mathcal{N}^*_h(\mathbf{w}(t))$, is also compact. The compactness of $\mathcal{N}^*_h(\mathbf{w}(t))$ combined with Eq. (66) implies

$$\lim_{\epsilon\downarrow 0}\lim_{j\to\infty}\sum_{h\in\mathcal{H}}\dot{\mathbf{b}}_h(t)\,\overline{\Lambda}_h(t,\epsilon,j) \le \sum_{h\in\mathcal{H}}\dot{\mathbf{b}}_h(t)\,\mathcal{N}^*_h(\mathbf{w}(t)) \le (1+\ln^{-1}(1/\delta))\sum_{h\in\mathcal{H}}\pi^P_h\mathcal{N}^*_h(\mathbf{w}(t)), \tag{67}$$

where the last inequality follows from the fact that $\dot{\mathbf{b}}_h(t) \leq \pi_h^P \frac{mq^P}{n^P} \leq \pi_h^P(1+\ln^{-1}(1/\delta))$ (Eq. (21)). Eqs. (65) and (67) together imply that

$$(\dot{\mathbf{a}}_k(t))_{k=1,\ldots,c_\mathcal{K}} \leq (1+\ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P \mathcal{N}_h^*(\mathbf{w}(t)). \tag{68}$$

We have verified that the all conditions in Definition 7.3 are met, and $\mathbf{w}(\cdot)$ is a fluid solution. This completes the proof of Proposition 7.8. □

7.4 Drift Properties of Fluid Solutions

We show in this subsection that the fluid solutions exhibit a certain contraction property with respect to the Lyapunov function, $L : \mathbb{R}_+^{c_\mathcal{K}+1} \to \mathbb{R}_+$, defined as:

$$L(\mathbf{w}) = \|\mathbf{w}\|_2 = \sqrt{\sum_{k=0}^{c_\mathcal{K}} \mathbf{w}_k^2}, \quad \mathbf{w} \in \mathbb{R}_+^{c_\mathcal{K}+1}. \tag{69}$$

The following proposition is the main result of this subsection, which shows that when m is sufficiently large, the value of the Lyapunov function always decreases by a constant amount starting from any initial condition with unit value.

Proposition 7.11 *Let m_F^* be the optimal value of* FLP*. Suppose that $mq^P > n^P$, and*

$$mq^A > \left(1+\frac{\ln(2c_\mathcal{H})+g_\delta}{\ln(1/\delta)}\right)\left(1+\frac{2c_\mathcal{H}^2 c_\mathcal{K}\bar{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1+\ln^{-1}(1/\delta))m_F^*. \tag{70}$$

Then, there exist $\tau, \epsilon' > 0$, such that, for any $\mathbf{w}^0 \in \mathbb{R}_+^{c_\mathcal{K}+1}$ with $L(\mathbf{w}^0) = 1$,

$$L(\mathbf{w}(\tau)) \leq 1-\epsilon', \quad \forall \mathbf{w} \in \mathcal{W}(\mathbf{w}^0). \tag{71}$$

The proof of the proposition is given in Appendix A.9. A main step of the proof is to couple the drift properties of the fluid solutions, a result of the workload creation in the Adaptive stage, to the structure of the Fundamental Linear Program (FLP) in Definition 6.1, and show that the contraction property holds as long as the system size is approximately greater than the optimal solution of FLP. To this end, we leverage an upper bound on the error of the coarse label estimate produced by the Preparation stage, and show that its accuracy is sufficiently high so that the appropriation of resouces in the Adaptive stage resembles an optimal solution to FLP.

7.5 Proof of Stability of Preparation and Adaptive Stages

We now complete the proof of Theorem 7.1 by establishing the joint stability of the Preparation and Adaptive stages. We will use the following version of the Foster-Lyapunov criterion. The main steps in this subsection are similar to those used in Massoulié (2007).

Proposition 7.12 (Theorem 8.13, Robert (2003)) *Let $\{X(t)\}_{t\in\mathbb{R}_+}$ be a Markov jump process on a countable state space, $\overline{\mathcal{X}}$. Suppose that there exists a function $L : \overline{\mathcal{X}} \to \mathbb{R}_+$, constants $C, \epsilon > 0$, and an integrable stopping time $\hat{\tau} > 0$, such that for all $x \in \overline{\mathcal{X}}$ such that $L(x) > C$, we have that*

$$\mathbb{E}\left(L(X(\hat{\tau})) \,\middle|\, X(0) = x\right) \leq L(x) - \epsilon\mathbb{E}\left(\hat{\tau} \,\middle|\, X(0) = x\right). \tag{72}$$

Suppose, in addition, that the set $\{x : L(x) \leq C\}$ is finite, and that $\mathbb{E}\left(L(X(\hat{\tau})) \,\middle|\, X(0) = x\right) < \infty$ for all $x \in \overline{\mathcal{X}}$. Then, $\{X(t)\}_{t\in\mathbb{R}_+}$ is positive recurrent.

Let $L(\cdot)$ be defined as in Eq. (69). Denote by $\{\mathbf{U}(t)\}_{t\in\mathbb{R}_+}$ the Markov process that describes the system dynamics in the Preparation and Adaptive stages, where the states are defined in Section 7.2, and by $\mathcal{U}$ the state space. We will denote by $\mathbb{P}_\mathbf{u}(\cdot)$ the probability distribution associated with the process $\mathbf{U}(\cdot)$ with initial condition $\mathbf{u} \in \mathcal{U}$. Define the function $\hat{L} : \mathcal{U} \to \mathbb{R}_+$

as the extension of $L(\cdot)$ on $\mathcal{U}$, i.e., $\hat{L}(\mathbf{U}(t)) = L(\mathbf{W}(t)) = \|\mathbf{W}(t)\|_2$. Let τ and ϵ' be defined as in Proposition 7.11. For every $\tilde{\mathbf{u}} \in \mathcal{U}$, let

$$\hat{\tau} = \hat{L}(\tilde{\mathbf{u}})\tau = L(\tilde{\mathbf{w}})\tau. \tag{73}$$

Consider the family of probability measures

$$\left\{\mathbb{P}_{\tilde{\mathbf{u}}}\big(\mathbf{W}_k(\hat{\tau})/\hat{L}(\tilde{\mathbf{u}}) \in \cdot\big)\right\}_{k=0,\dots,c_{\mathcal{K}},\tilde{\mathbf{u}}\in\mathcal{U},\tilde{\mathbf{w}}\neq 0}. \tag{74}$$

We first show that the above family is uniformly integrable. We have the dominance relation:

$$\mathbf{W}_k(t) \le \mathbf{W}_k(0) + \max\{v_\delta, n^P\}\Xi_0(t), \quad t \in \mathbb{R}_+, k = 0, \dots, c_{\mathcal{K}}. \tag{75}$$

To see why this is true, recall that a job creates n^P inspections in the Preparation stage ($k = 0$) and at most v_δ total inspections in the Adaptive stage ($k = 1, \dots, c_{\mathcal{K}}$). Hence, the second term on the right-hand side of Eq. (75) dominates the total number of inspections that could have been added to $\mathbf{W}_k(\cdot)$ by time t. Scaling both sides of Eq. (75) by $\hat{L}(\tilde{\mathbf{u}})$ and setting $t = \hat{\tau}$, we have that, when $\mathbf{U}(0) = \tilde{\mathbf{u}}$,

$$\begin{aligned}
\mathbf{W}_k(\hat{\tau})/\hat{L}(\tilde{\mathbf{u}}) \le& \tilde{\mathbf{w}}_k/\hat{L}(\tilde{\mathbf{u}}) + \max\{v_\delta, n^P\}\Xi_0(\hat{\tau})/\hat{L}(\tilde{\mathbf{u}}) \\
=& \tilde{\mathbf{w}}_k/L(\tilde{\mathbf{w}}) + \max\{v_\delta, n^P\}\Xi_0(\hat{L}(\tilde{\mathbf{u}})\tau)/\hat{L}(\tilde{\mathbf{u}}) \\
\overset{(a)}{\le}& 1/\alpha_1 + \max\{v_\delta, n^P\}\Xi_0(\hat{L}(\tilde{\mathbf{u}})\tau)/\hat{L}(\tilde{\mathbf{u}}),
\end{aligned} \tag{76}$$

for some constant $\alpha_1 > 0$, where step (a) follows from the first inequality in Eq. (237) of Lemma B.1 in Appendix B. Since $\Xi_0(\cdot)$ is a unit-rate Poisson process, we have $\mathbb{E}\left(\Xi_0(\hat{L}(\tilde{\mathbf{u}})\tau)/\hat{L}(\tilde{\mathbf{u}})\right) = \mathrm{Var}\left(\Xi_0(\hat{L}(\tilde{\mathbf{u}})\tau)/\hat{L}(\tilde{\mathbf{u}})\right) = \tau$. The above arguments show that the first and second moments of $\mathbf{W}_k(\hat{\tau})/\hat{L}(\tilde{\mathbf{u}})$ are both bounded from above uniformly over all $\tilde{\mathbf{u}} \in \mathcal{U}$, so long as $\tilde{\mathbf{w}} \neq 0$. This implies the uniform integrability of the family of distributions in Eq. (74). Using again Eq. (237) of Lemma B.1, we have that for some constant $\alpha_2 > 0$

$$\hat{L}(\mathbf{U}(\hat{\tau}))/\hat{L}(\tilde{\mathbf{u}}) = L(\mathbf{W}(\hat{\tau}))/\hat{L}(\tilde{\mathbf{u}}) \le \frac{\alpha_2\|\mathbf{W}(\hat{\tau})\|_\infty}{\hat{L}(\tilde{\mathbf{u}})} \le \frac{\alpha_2}{\alpha_1} + \alpha_2 \max\{v_\delta, n^P\}\Xi_0(\hat{L}(\tilde{\mathbf{u}})\tau)/\hat{L}(\tilde{\mathbf{u}}). \tag{77}$$

Using the same arguments as those following Eq. (76), we have that the set of probability measures

$$\left\{\mathbb{P}_{\tilde{\mathbf{u}}}\big(\hat{L}(\mathbf{U}(\hat{\tau}))/\hat{L}(\tilde{\mathbf{u}}) \in \cdot\big)\right\}_{\tilde{\mathbf{u}}\in\mathcal{U},\tilde{\mathbf{w}}\neq 0} \tag{78}$$

is also uniformly integrable.

We now prove the validity of Eq. (72) in our context. In particular, we will show that there exist $A, \epsilon > 0$, such that for all $\tilde{\mathbf{u}} \in \mathcal{U}$, if $\hat{L}(\tilde{\mathbf{u}}) \ge A$, then we have that

$$\mathbb{E}_{\tilde{\mathbf{u}}}\left(\hat{L}(\mathbf{U}(\hat{L}(\tilde{\mathbf{u}})\tau))\right) \le \hat{L}(\tilde{\mathbf{u}})(1 - \epsilon\tau), \tag{79}$$

where we have used the substitution $\hat{\tau} = \hat{L}(\tilde{\mathbf{u}})\tau$. By the definitions of $\hat{L}$ and $\hat{\tau}$, we have that

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbf{u}}}\left(\hat{L}(\mathbf{U}(\hat{L}(\tilde{\mathbf{u}})\tau))/\hat{L}(\mathbf{u})\right) =& \mathbb{E}_{\tilde{\mathbf{u}}}\big(L(\mathbf{W}(L(\tilde{\mathbf{w}})\tau))/L(\tilde{\mathbf{w}})\big) \overset{(a)}{=} \mathbb{E}_{\tilde{u}}\big(L(L(\tilde{\mathbf{w}})^{-1}\mathbf{W}(L(\tilde{\mathbf{w}})\tau))\big) \\
=& \mathbb{E}_{\tilde{u}}\big(L(\mathbf{W}(L(\tilde{\mathbf{w}}), \tau))\big),
\end{aligned} \tag{80}$$

where step (a) is based on the second property in Lemma B.1 (one-homogeneity of $L(\cdot)$). In light of Eq. (80), the condition in Eq. (79) is equivalent to

$$\mathbb{E}_{\tilde{\mathbf{u}}}\left(L(\mathbf{W}(L(\tilde{\mathbf{w}}), \tau))\right) \le 1 - \epsilon\tau. \tag{81}$$

We will now prove Eq. (81) by contradiction. For the sake of contradiction, suppose that there exists a sequence of initial conditions, $\{\tilde{\mathbf{u}}^s\}_{s\in\mathbb{N}}$, where $\hat{L}(\tilde{\mathbf{u}}^s) \to \infty$ as $s \to \infty$, such that

$$\limsup_{s\to\infty} \mathbb{E}_{\tilde{\mathbf{u}}^s}\left(L(\mathbf{W}^s(L(\tilde{\mathbf{w}}^s),\tau))\right) \geq 1, \tag{82}$$

where we have re-introduced the superscript s in $\mathbf{W}$ to signify the initial condition $\mathbf{W}^s(0) = \tilde{\mathbf{w}}^s$. In particular, after scaling, we can define the quantity $\mathbf{w}^s$:

$$\mathbf{w}^s \stackrel{\triangle}{=} \mathbf{W}^s(L(\tilde{\mathbf{w}}^s),0) = \mathbf{W}^s(0)/L(\tilde{\mathbf{w}}^s) = \tilde{\mathbf{w}}^s/L(\tilde{\mathbf{w}}^s). \tag{83}$$

Since $L(\mathbf{w}) = \|\mathbf{w}\|_2$, we have that $\|\mathbf{w}^s\|_2 = L(\mathbf{w}^s/L(\tilde{\mathbf{w}}^s)) = 1$, i.e., $\mathbf{w}^s$ belongs to the compact set $\{\mathbf{w} \in \mathbb{R}_+^{c_\mathcal{K}+1} : \|\mathbf{w}\|_2 = 1\}$. The compactness implies that there exists $\mathbf{w}^0 \in \mathbb{R}_+^{c_\mathcal{K}+1}$, with $L(\mathbf{w}^0) = \|\mathbf{w}^0\|_2 = 1$, and a sub-sequence of $\{\mathbf{w}^s\}$, $\{\mathbf{w}^{s_i}\}_{i\in\mathbb{N}}$, such that

$$\lim_{i\to\infty} \mathbf{w}^{s_i} = \mathbf{w}^0. \tag{84}$$

Since $L(\tilde{\mathbf{w}}^s) = \hat{L}(\tilde{\mathbf{u}}^s) \to \infty$ as $s \to \infty$ by our assumption, by Proposition 7.4 and Eq. (84), we have

$$\lim_{i\to\infty} \inf_{\mathbf{y}\in\mathcal{V}(\mathbf{w}^0,\tau)} \|\mathbf{y} - \mathbf{W}^{s_i}(L(\tilde{\mathbf{w}}^{s_i}),\tau)\|_2 = 0, \quad \text{a.s.}, \tag{85}$$

where $\mathcal{V}(\mathbf{w}^0,t)$ is defined to be the set of all states of the fluid solution at time $t \in \mathbb{R}_+$, starting with an initial condition $\mathbf{w}^0$: $\mathcal{V}(\mathbf{w}^0,t) = \bigcup_{\mathbf{w}(\cdot)\in\mathcal{W}(\mathbf{w}^0)} \mathbf{w}(t)$. It can be verified from the definition of the fluid solutions that the set $\mathcal{V}(\mathbf{w}^0,t)$ is compact for all t and $\mathbf{w}^0$. The compactness of $\mathcal{V}(\mathbf{w}^0,\tau)$ along with Eq. (85) implies that there exist $\bar{\mathbf{y}} \in \mathcal{V}(\mathbf{w}^0,\tau)$ and a sub-sequence of $\{\tilde{\mathbf{w}}^{s_i}\}$, $\{\tilde{\mathbf{w}}^{s_{i_j}}\}_{j\in\mathbb{N}}$, such that

$$\lim_{j\to\infty} \mathbf{W}^{s_{i_j}}(L(\tilde{\mathbf{w}}^{s_{i_j}}),\tau) = \bar{\mathbf{y}}, \quad \text{a.s.} \tag{86}$$

Since $L(\cdot)$ is a continuous function, the above equation further implies that

$$\lim_{j\to\infty} L(\mathbf{W}^{s_{i_j}}(L(\tilde{\mathbf{w}}^{s_{i_j}}),\tau)) = L(\bar{\mathbf{y}}) \leq 1-\epsilon', \quad \text{a.s.}, \tag{87}$$

where the last inequality follows from the fact that $L(\mathbf{w}^0) = 1$, $\bar{\mathbf{y}} \in \mathcal{V}(\mathbf{w}^0,\tau)$, and Proposition 7.11. As was shown earlier, the family of distributions, $\left\{\mathbb{P}_{\tilde{\mathbf{u}}}\left(\hat{L}(\mathbf{U}(\hat{\tau}))/\hat{L}(\tilde{\mathbf{u}}) \in \cdot\right)\right\}_{\tilde{\mathbf{u}}\in\mathcal{U},\tilde{\mathbf{w}}\neq 0}$, is uniformly integrable (Eq. (78)), and hence one of its subsets

$$\left\{\mathbb{P}_{\tilde{\mathbf{u}}^{s_{i_j}}}\left(L(\mathbf{W}^{s_{i_j}}(L(\tilde{\mathbf{w}}^{s_{i_j}}),\tau) \in \cdot\right)\right\}_{i\in\mathbb{N}} \tag{88}$$

is also uniformly integrable. Combining Eq. (87) with the above uniform integrability implies that

$$\lim_{j\to\infty} \mathbb{E}_{\tilde{\mathbf{u}}^{s_{i_j}}}\left(L(\mathbf{W}^{s_{i_j}}(L(\tilde{\mathbf{w}}^{s_{i_j}}),\tau))\right) = L(\bar{\mathbf{y}}) \leq 1-\epsilon' < 1, \tag{89}$$

which contradicts with Eq. (82). This completes the proof of Eq. (79).

The fact that $\mathbb{E}_{\tilde{\mathbf{u}}}(\tilde{L}(\mathbf{U}(\hat{\tau})) < \infty$ for all $\tilde{\mathbf{u}}$ is readily verifiable from Eq. (77). Finally, for any finite C, the set $\{\mathbf{u} \in \mathcal{U} : L(\mathbf{u}) < C\}$ contains states where the total number of jobs currently in system is bounded. Since each job can receive at most $n^P + v_\delta$ inspections, and the result of each inspection belongs to a finite set, it follows (Section 7.2) that the set $\{\mathbf{u} \in \mathcal{U} : L(\mathbf{u}) < C\}$ is finite. We have thus verified all the conditions in Proposition 7.12, and established the positive recurrence of $\mathbf{U}(\cdot)$. This completes the proof of Theorem 7.1. □

8 The Residual Stage

We now turn to the residual stage. Assuming the stability of the first two stages, the following result gives a sufficient condition for system stability, expressed in term of the total processing resources devoted to the Residual stage; the proof is given in Appendix A.10.

Proposition 8.1 *Suppose that the Preparation and Adaptive stages are stable. Then, the overall system is stable whenever*

$$mq^R > 3c_\mathcal{H}\zeta_0(1 + \ln(4c_\mathcal{H})\ln^{-1}(1/\delta)). \tag{90}$$

9 Proof of Theorem 3.1

We now complete the proof of our main result, Theorem 3.1. We begin by showing that the policy proposed in Section 5 is δ-accurate: for *every* job i that departs from the system, the probability of it being misclassified is at most δ. The proof is given in Appendix A.11.

Proposition 9.1 *Fix $\delta \in (0,1)$. Under the inspection policy described in Section 5,*

$$\mathbb{P}\left(\widehat{H}_i \neq H_i \,\middle|\, H_i = h\right) \leq \delta, \quad \forall h \in \mathcal{H}, i \in \mathbb{N}. \tag{91}$$

Completing the Proof of Theorem 3.1. We now show that the minimum value of m required to stabilize the system under the proposed inspection policy satisfies the inequality in Eq. (11). Without loss of generality, we may assume that $m \geq m_F^*$, where m_F^* is the optimal value of FLP (Definition 6.1). Combining Theorem 7.1 and Proposition 8.1, we have that the system is stable as long as all of the following hold:

$$mq^A > \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \frac{2c_{\mathcal{H}}^2 c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1 + \ln^{-1}(1/\delta))m_F^*, \tag{92}$$

$$mq^P > \zeta_0 \ln\ln(1/\delta), \quad mq^R > 3c_{\mathcal{H}}\zeta_0(1 + \ln(4c_{\mathcal{H}})\ln^{-1}(1/\delta)), \tag{93}$$

Recall from the definition of our policy in Eqs. (21) and (22) that $q^P = \left(\zeta_0 \ln\ln(1/\delta) + \ln^{-1}(1/\delta)\right)/m$, $q^R = \left[3c_{\mathcal{H}}\zeta_0(1 + \ln(4c_{\mathcal{H}})\ln^{-1}(1/\delta)) + 1\right]/m$, and $q^A = 1 - q^P - q^R$. Because $m \geq m_F^* \geq \overline{d}^{-1}\ln(1/\delta)$ (Lemma 6.3), it is not difficult to verify that q^P, q^R and q^A lie in $(0,1)$ for all sufficiently small δ. Note that that Eq. (93) is automatically satisfied by construction, whenever $q^P, q^R \in (0,1)$. Therefore, the system is stable if Eq. (92) is true, which, by inserting the expressions for q^P and q^R, requires that

$$m > \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \frac{2c_{\mathcal{H}}^2 c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1 + \ln^{-1}(1/\delta))m_F^* + \iota_\delta, \tag{94}$$

where $\iota_\delta \triangleq \zeta_0 \ln\ln(1/\delta) + 3c_{\mathcal{H}}\zeta_0(1 + \ln(4c_{\mathcal{H}})\ln^{-1}(1/\delta)) + 1 + \ln^{-1}(1/\delta)$. By dividing both sides of Eq. (94) by m_F^*, the condition becomes

$$\frac{m}{m_F^*} > \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \frac{2c_{\mathcal{H}}^2 c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1 + \ln^{-1}(1/\delta)) + \frac{\iota_\delta}{m_F^*}. \tag{95}$$

The two terms on the right-hand side of the above equation can be bounded as follows. For the first term, recall from the definition in Eq. (29) that $g_\delta = 3\overline{z}\sqrt{c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\ln\ln(1/\delta)}$. It is not difficult to see that there exist constants $\bar{c}_1$ and $\bar{\delta}_1 > 0$, independent of π, such that for all $\delta < \bar{\delta}_1$

$$\begin{aligned}&\left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \frac{2c_{\mathcal{H}}^2 c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}\ln(1/\delta)}\right)(1 + \ln^{-1}(1/\delta)) \\ <&1 + \frac{\bar{c}_1\sqrt{\ln(1/\delta)\ln\ln(1/\delta)}}{\ln(1/\delta)} = 1 + \bar{c}_1\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}.\end{aligned} \tag{96}$$

For the second term in Eq. (95), ι_δ/m_F^*, note that the dominating term in ι_δ is of order $\ln\ln(1/\delta)$. Recall from Lemma 6.3 that $m_F^* \geq \overline{d}^{-1}\ln(1/\delta)$, and hence there exist constants $\bar{c}_2$ and $\bar{\delta}_2 > 0$, independent of π, such that for all $\delta < \bar{\delta}_2$, we have that $\frac{\iota_\delta}{m_F^*} \leq \frac{\iota_\delta}{\overline{d}^{-1}\ln(1/\delta)} \leq \bar{c}_2\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}$. This, along with Eqs. (95) and (96), implies that, for all $\delta < \min\{\bar{\delta}_1, \bar{\delta}_2\}$, it suffices to have

$$\frac{m}{m_F^*} > 1 + \bar{c}_1\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}} + \bar{c}_2\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}. \tag{97}$$

Finally, the lower bound in Proposition 6.2 shows that the optimal system size, $m^*(\delta,\pi)$, satisfies

$$m^*(\delta,\pi) \geq (1-\delta)\left[1-\left(\ln\frac{1}{1-\delta}+e^{-1}\right)\ln^{-1}(1/\delta)\right]m_F^*, \tag{98}$$

for all sufficiently small δ. In particular, there exist constants $\bar{c}_3$ and $\bar{\delta}_3 > 0$, independent of π, such that for all $\delta < \bar{\delta}_3$

$$m_F^* \leq \left(1+\frac{\bar{c}_3}{\ln(1/\delta)}\right)m^*(\delta,\pi). \tag{99}$$

Substituting Eq. (99) into Eq. (97), and noticing that the leading term in $\left(\frac{m}{m^*(\delta,\pi)}-1\right)$ is of order $\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}$, we conclude that there exist $c_0, \delta_0 > 0$, independent of π, such that, for all $\delta \in (0,\delta_0)$, the system is stable under the proposed policy whenever

$$\frac{m}{m^*(\delta,\pi)} > 1 + c_0\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}. \tag{100}$$

Together with Proposition 9.1, this completes the proof of Theorem 3.1. □

10 Concluding Remarks

The main objective of this paper is to understand the design principles and fundamental limitations involved as one aims to efficiently classify a large set of items using a finite amount of processing resources. The main result demonstrates a prior-oblivious inspection architecture that asymptotically uses the minimum number of experts, in the regime where the required classification error tends to zero. More broadly speaking, our result could be viewed as an attempt towards understanding how to build effective processing architectures and algorithms for large-scale statistical learning or information extraction tasks, given limited resources or processing power. We believe that there are many other problems in this domain, situated at the intersection between stochastic modeling and statistics, that may be of interest for future research.

Our model relies on a few assumptions which we hope can be relaxed in future work. For instance, we assume that the outcome distributions are perfectly known to the decision maker, whereas in reality they may be initially unknown and must be learned. We also assume that the duration of an inspection depends only on the expert type and not on the nature of the job. Removing these assumptions appears challenging and would likely require new methodologies or formulations.

The present paper leaves open a few interesting questions. First, we have thus far required the inspection policies to be stable, while the more refined metric of delays experienced by the jobs has not been investigated. Second, the resource efficiency property of the proposed policy applies only in the regime where the allowable error is close to zero. If significantly greater errors can be tolerated, however, it becomes less clear what inspection policy one should choose and whether a different criterion of resource efficiency should be adopted. Finally, it would be interesting to identify, and provide rigorous guarantees for, simpler heuristic policies, such as the one we propose in Appendix D, which are more easily implementable in practical applications.

References

Alizamir, S., de Véricourt, F., and Sun, P. (2013). Diagnostic accuracy under congestion. *Management Science*, 59(1):157–171.

Baccelli, F. and Brémaud, P. (2003). *Elements of queueing theory: Palm Martingale calculus and stochastic recurrences.* Springer, second edition.

Baker, K. and von Beers, J. (1996). Shmoo plotting: The black art of IC testing. In *Proceedings of IEEE International Test Conference*, pages 932–933.

Berge, C. (1963). *Topological Spaces: Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity*. Courier Corporation.

Bernstein, M. S., Karger, D. R., Miller, R. C., and Brandt, J. (2012). Analytic methods for optimizing realtime crowdsourcing. In *Proceedings of Collective Intelligence*.

Bimpikis, K. and Markakis, M. G. (2016). Learning and hierarchies in service systems. *under submission*.

Bramson, M. (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems*, 30(1-2):89–140.

Chernoff, H. (1959). Sequential design of experiments. *Annals of Mathematical Statistics*, 30(3):755–770.

Dai, J. G. (1995). On positive harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *The Annals of Applied Probability*, pages 49–77.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In *29th Annual Conference on Learning Theory*, pages 998–1027.

Gerdtz, M. F. and Bucknall, T. K. (2001). Triage nurses' clinical decision making. An observational study of urgency assessment. *Journal of Advanced Nursing*, 35(4):550–561.

Grimmett, G. and Stirzaker, D. (2001). *Probability and random processes*. Oxford university press.

Harrison, J. M. and López, M. J. (1999). Heavy traffic resource pooling in parallel-server systems. *Queueing systems*, 33(4):339–368.

Harrison, J. M. and Sunar, N. (2015). Investment timing with incomplete information and multiple means of learning. *Operations Research*, 62(2):442–457.

Ho, C., Jabbari, S., and Vaughan, J. W. (2013). Adaptive task assignment for crowdsourced classification. In *Proceedings of the International Conference on Machine Learning*.

Johari, R., Kamble, V., and Kanoria, Y. (2016). Know your customer: Multi-armed bandits with capacity constraints. *unpublished manuscript*.

Karger, D. R., Oh, S., and Shah, D. (2014). Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research*, 62(1):1–24.

Massoulié, L. (2007). Structural properties of proportional fairness: stability and insensitivity. *Annals of Applied Probability*, 17(3):809–839.

Retsef, L., Magnanti, T., and Shaposhnik, Y. (2015). Scheduling with testing. *under submission*.

Robert, P. (2003). *Stochastic Networks and Queues*. Springer.

Rybko, A. N. and Stolyar, A. (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. *Problemy Peredachi Informatsii*, 28(3):3–26.

Siegmund, D. (2013). *Sequential analysis: tests and confidence intervals*. Springer Science & Business Media.

Talreja, R. and Whitt, W. (2008). Fluid models for overloaded multiclass many-server queueing systems with first-come, first-served routing. *Management Science*, 54(8):1513–1527.

Tassiulas, L. and Ephremides, A. (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1948.

Tsitsiklis, J. N. and Xu, K. (2012). On the power of (even a little) resource pooling. *Stochastic Systems*, 2(1):1–66.

Tsitsiklis, J. N. and Xu, K. (2017). Flexible queueing architectures. to appear in *Operations Research*. https://doi.org/10.1287/opre.2017.1620.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16(2):117–186.

Ye, H.-Q., Ou, J., and Yuan, X.-M. (2005). Stability of data networks: Stationary and bursty models. *Operations Research*, 53(1):107–125.

Electronic Companion: On the Capacity of Information Processing Systems [8]

Laurent Massoulié
Microsoft Research-Inria Joint Centre
Palaiseau, France
laurent.massoulie@inria.fr

Kuang Xu
Graduate School of Business
Stanford University
kuangxu@stanford.edu

A Proofs

A.1 Proof of Lemma 4.2

Proof. Fix $i \in \mathbb{N}$. Denote by N_i the total number of inspections job i receives before departing from the system. Let $Y_{i,n}$ be the inspection history of job i up till the nth inspection:

$$Y_{i,n} = \{(K_{i,j}, X_{i,j})\}_{j=1,\dots,n}. \tag{101}$$

We will use $y_n = \{(k_j, x_j)\}_{j=1,\dots,n}$ to denote a specific realization of $Y_{i,n}$, with $y_0 \overset{\triangle}{=} \emptyset$. Define $q_h(y_n)$ as the probability

$$q_h(y_n) = \mathbb{P}(N_i = n, Y_{i,n} = y_n \,\big|\, H_i = h), \quad n \in \mathbb{Z}_+, \tag{102}$$

and $p_h(y_n)$ as the likelihood associated with an inspection history, $y_n = \{(k_j, x_j)\}_{j=1,\dots,n}$ under label h:

$$p_h(y_n) = \prod_{j=1}^{n} p(h, k_j, x_j), \quad n \in \mathbb{N}, \tag{103}$$

with $p_h(y_0) \overset{\triangle}{=} 1$. We have that, for all $n \in \mathbb{Z}_+$,

$$\begin{aligned} q_h(y_n) =& \mathbb{P}(N_i = n, Y_{i,n} = y_n \,\big|\, H_i = h) \\ =& \mathbb{P}(N_i = n \,\big|\, Y_{i,n} = y_n) \prod_{j=1}^{n} p(h, k_j, x_j) \mathbb{P}(K_{i,j} = k_j, N_i \geq j \,\big|\, Y_{i,j-1} = y_{j-1}) \\ =& \Big(\prod_{j=1}^{n} p(h, k_j, x_j) \Big) \Big(\mathbb{P}(N_i = n \,\big|\, Y_{i,n} = y_n) \prod_{j=1}^{n} \mathbb{P}(K_{i,j} = k_j, N_i \geq j \,\big|\, Y_{i,j-1} = y_{j-1}) \Big) \\ =& p_h(y_n) v(y_n) \end{aligned} \tag{104}$$

where y_j, $1 \leq j \leq n$, denotes the first j inspections in y_n, and $v(y_n) \overset{\triangle}{=} \mathbb{P}(N_i = n \,\big|\, Y_{i,n} = y_n) \prod_{j=1}^{n} \mathbb{P}(K_{i,j} = k_j, N_i \geq j \,\big|\, Y_{i,j-1} = y_{j-1})$, for $n \in \mathbb{N}$, and $v(y_0) \overset{\triangle}{=} \mathbb{P}(N_i = 0)$. Importantly, note that because the inspection policy does not have access to the true label of a job, h, $v(y_n)$ only can only depend on the realization of the history, and not on the value of h.

We have that

$$\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})} \,\Big|\, \widehat{H}_i = h, H_i = h \right) = \mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})} \mathbb{I}(\widehat{H}_i = h) \,\Big|\, H_i = h \right) / \mathbb{P}(\widehat{H}_i = h \,\big|\, H_i = h), \tag{105}$$

[8] March 2016; revised February 2017.

where

$$
\begin{aligned}
\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\mathbb{I}(\widehat{H}_i = h) \,\Big|\, H_i = h\right) &\overset{(a)}{=} \sum_{n=0}^{\infty}\sum_{y_n}\frac{p_l(y_n)}{p_h(y_n)}\mathbb{P}(N_i = n, Y_{i,n} = y_n, \widehat{H}_i = h \,|\, H_i = h) \\
&= \sum_{n=0}^{\infty}\sum_{y_n}\frac{p_l(y_n)}{p_h(y_n)}q_h(y_n)\mathbb{P}(\widehat{H}_i = h \,|\, N_i = n, Y_{i,n} = y_n, H_i = h) \\
&\overset{(b)}{=} \sum_{n=0}^{\infty}\sum_{y_n} p_l(y_n)v(y_n)\mathbb{P}(\widehat{H}_i = h \,|\, N_i = n, Y_{i,n} = y_n, H_i = h) \\
&\overset{(c)}{=} \sum_{n=0}^{\infty}\sum_{y_n} p_l(y_n)v(y_n)\mathbb{P}(\widehat{H}_i = h \,|\, N_i = n, Y_{i,n} = y_n, H_i = l) \\
&\overset{(d)}{=} \sum_{n=0}^{\infty}\sum_{y_n} q_l(y_n)\mathbb{P}(\widehat{H}_i = h \,|\, N_i = n, Y_{i,n} = y_n, H_i = l) \\
&= \sum_{n=0}^{\infty}\sum_{y_n} \mathbb{P}(N_i = n, Y_{i,n} = y_n, \widehat{H}_i = h \,|\, H_i = l) \\
&= \mathbb{P}(\widehat{H}_i = h \,|\, H_i = l). \qquad (106)
\end{aligned}
$$

When $n \geq 1$, the second summation in step (a) is over all possible realizations of $Y_{i,n}$ for which $p_h(y_n)$ is non-zero, and when $n = 0$, the summation is over the single element, $\emptyset$. Steps (b) and (d) follow from Eq. (104). For step (c), we used the fact that conditioning on the event $\{N_i = n, Y_{i,n} = y_n\}$, the event $\{\widehat{H}_i = h\}$ is independent from the true label of job i. Eqs. (105) and (106) combined yield that

$$
\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})} \,\Big|\, \widehat{H}_i = h, H_i = h\right) = \frac{\mathbb{P}(\widehat{H}_i = h \,|\, H_i = l)}{\mathbb{P}(\widehat{H}_i = h \,|\, H_i = h)} \leq \frac{\delta}{1-\delta}, \qquad (107)
$$

where the last equality follows from the assumption of the policy being δ-accurate.

Define

$$
\delta_h = \mathbb{P}(\widehat{H}_i \neq h \,|\, H_i = h). \qquad (108)
$$

Using essentially the same steps of deduction as those leading to Eq. (107), we have that

$$
\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})} \,\Big|\, \widehat{H}_i \neq h, H_i = h\right) = \frac{\mathbb{P}(\widehat{H}_i \neq h \,|\, H_i = l)}{\mathbb{P}(\widehat{H}_i \neq h \,|\, H_i = h)} \leq \frac{1}{\delta_h} \qquad (109)
$$

We are now ready to prove the main claim of the lemma. Fix $\delta \in (0, 1/2)$. We have that

$$
\begin{aligned}
\mathbb{E}\left(S_i^F(h,l)\,\middle|\,H_i=h\right) =&\mathbb{E}\left(\sum_{j=1}^{N_i}\ln\frac{p(h,K_{i,j},X_{i,j})}{p(l,K_{i,j},X_{i,j})}\,\middle|\,H_i=h\right)\\
=&\mathbb{E}\left(-\ln\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\,\middle|\,H_i=h\right)\\
=&-\mathbb{E}\left(\ln\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\,\middle|\,\widehat{H}_i=h,H_i=h\right)\mathbb{P}(\widehat{H}_i=h\,|\,H_i=h)\\
&-\mathbb{E}\left(\ln\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\,\middle|\,\widehat{H}_i\neq h,H_i=h\right)\mathbb{P}(\widehat{H}_i\neq h\,|\,H_i=h)\\
\overset{(a)}{\geq}&-\ln\left(\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\,\middle|\,\widehat{H}_i=h,H_i=h\right)\right)\mathbb{P}(\widehat{H}_i=h\,|\,H_i=h)\\
&-\ln\left(\mathbb{E}\left(\frac{p_l(Y_{i,N_i})}{p_h(Y_{i,N_i})}\,\middle|\,\widehat{H}_i\neq h,H_i=h\right)\right)\mathbb{P}(\widehat{H}_i\neq h\,|\,H_i=h)\\
\overset{(b)}{\geq}&(1-\delta)\ln\frac{1-\delta}{\delta}+\delta_h\ln\delta_h\\
\overset{(c)}{\geq}&(1-\delta)\ln\frac{1-\delta}{\delta}-e^{-1}.
\end{aligned}
\tag{110}
$$

Step (a) is based on Jensen's inequality, by noting that that $-\ln(\cdot)$ is a convex function. Step (b) follows from Eqs. (107) and (109), and the policy being δ-accurate. Step (c) is based on the observation that $(x\ln x)$ is monotonically decreasing over $x\in(0,e^{-1})$ and monotonically increasing over $(e^{-1},1)$. This proves Lemma 4.2. □

A.2 Proof of Proposition 6.2

Proof. Fix a stable inspection policy, ψ, and consider a time-stationary trajectory of all processes needed to describe the system and policy states. Denote by N_i the total number of inspections, and by $N_{i,k}$ the number of inspections from type-k experts, received by job i before a classification is produced.

Denote by $\{A(t)\}_{t\in\mathbb{R}_+}$ the Poisson point process associated with job arrivals to the system, $\{T_i^A\}_{i\in\mathbb{N}}$ the corresponding arrival times, and $\mathbb{P}_A^0$ the Palm probability associated with $A(\cdot)$ (intuitively, this can be thought of as the original probability law conditional on there being an arrival at time 0; see Baccelli and Brémaud (2003) for a detailed treatment). Note that the intensity of $A(\cdot)$, λ_A, is equal to 1 by assumption. We first show the following inequality:

$$
\mathbb{E}_A^0(N_{i,k})\leq mr_k. \tag{111}
$$

Note that by stationarity the left-hand side is independent of the arrival index, i, which we may then take to be equal to 1, i.e., by considering the arrival at time 0 under the Palm probability $\mathbb{P}_A^0$. To show Eq. (111), denote by $\{D_k(t)\}_{t\in\mathbb{R}_+}$ the point process corresponding to inspection completions by type-k experts, $\{T_j^k\}_{j\in\mathbb{N}}$ the corresponding completion times, and $\mathbb{P}_k^0$ the Palm probability associated with $D_k(\cdot)$. Let λ_k be the intensity of $\{D_k(t)\}_{t\in\mathbb{R}_+}$, i.e., the average number of inspection completions per unit of time. For each $j\in\mathbb{N}$, denote by $i_k(j)$ the index of the job for whom an inspection is completed at time T_j^k, and let F_j^k the difference between T_j^k and the arrival time of job $i_k(j)$, i.e.,

$$F_j^k = T_j^k - T_{i_k(j)}^A. \tag{112}$$

We then have

$$\begin{aligned}
\mathbb{E}_A^0\left(N_{i,k}\right) &\overset{(a)}{=} \frac{1}{\lambda_A}\mathbb{E}\left(\sum_{i\in\mathbb{N}} N_{i,k}\mathbf{1}_{T_i^A\in[0,1]}\right)\\
&\overset{(b)}{=}\mathbb{E}\left(\sum_{i\in\mathbb{N}} N_{i,k}\mathbf{1}_{T_i^A\in[0,1]}\right)\\
&\overset{(c)}{=}\mathbb{E}\left(\sum_{j\in\mathbb{N}} \mathbf{1}_{(T_j^k-F_j^k)\in[0,1]}\right)\\
&\overset{(d)}{=}\lambda_k\mathbb{E}_k^0\left(\int_{\mathbb{R}} \mathbf{1}_{(t-F_1^k)\in[0,1]}dt\right)\\
&\overset{(e)}{=}\lambda_k.
\end{aligned} \tag{113}$$

Step (a) follows from definition of Palm probability, (b) from $\lambda_A = 1$, and (c) from noticing that every inspection associated with the $N_{i,k}$ for jobs arriving during $[0,1]$, i.e., $\{i : T_i^A \in [0,1]\}$, corresponds to exactly one completion time, T_j^k, such that $(T_j^k - F_j^k) \in [0,1]$. Step (d) is an application of the Campbell-Little-Mecke formula (Eq. (3.3.3) of Baccelli and Brémaud (2003)), and (e) follows from the fact that the integral in (d) evaluates to 1, as is easily seen by a change of variables $u = t - F_1^k$. To obtain Eq. (111), it suffices to note that $\lambda_k \le mr_k$: the maximum number of inspections is obtained when the experts never idle, in which case the intensity of inspection completions is exactly mr_k.

Let us now condition on the type of the job arriving at time 0, H_i, and introduce the notation

$$n_{h,k} = \mathbb{E}_A^0(N_{i,k}|H_i = h), \quad h \in \mathcal{H}, k \in \mathcal{K}. \tag{114}$$

Recalling that the probability of H_i being h is π_h, we have that

$$\mathbb{E}_A^0(N_{i,k}) = \sum_{h\in\mathcal{H}} n_{h,k}\pi_h, \tag{115}$$

and Eq. (111) can be equivalently written as

$$\sum_{h\in\mathcal{H}} n_{h,k}\pi_h \le mr_k. \tag{116}$$

Recall that $S_i^F(h,l)$ is the cumulative log-likelihood ratio by the time when job i departs from the system:

$$S_i^F(h,l) = \sum_{j=1}^{N_i} Z_{i,j}(h,l,K_{i,j}). \tag{117}$$

We have that, for any job i and any fixed label h,

$$\begin{aligned}\sum_{k\in\mathcal{K}} n_{h,k} D(h,l,k) = &\sum_{k\in\mathcal{K}} \mathbb{E}^0_A(N_{i,k}|H_i = h) D(h,l,k) \\ \overset{(a)}{=}&\mathbb{E}^0_A\Big(\sum_{j=1}^{N_i} Z_{i,j}(h,l,K_{i,j}) \,\Big|\, H_i = h\Big) \\ =&\mathbb{E}^0_A(S^F_i(h,l) \,|\, H_i = h) \\ \overset{(b)}{\geq}&(1-\delta)\ln\frac{1-\delta}{\delta} - e^{-1}. \end{aligned} \tag{118}$$

Step (a) follows from Wald's identity, by noting that the responses from a given expert type on job i are $i.i.d$, and the responses are independent across different expert types (cf. Eq. (18) in Garivier and Kaufmann (2016)). Step (b) follows from Lemma 4.2, where we note that since the inspection outcomes are independent from the arrival times of the jobs, the lower bound derived in Lemma 4.2 continues to apply under the Palm probability.

From Eqs. (116) and (118), we conclude that the number of experts, m, under any stable, δ-accurate inspection policy must be no less than the optimal value of the following linear program:

$$\text{minimize} \quad m \tag{119}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}} n_{h,k} D(h,l,k) \geq (1-\delta)\ln\frac{1-\delta}{\delta} - e^{-1}, \quad \forall h,l \in \mathcal{H}, h\neq l, \tag{120}$$

$$\sum_{h\in\mathcal{H}} n_{h,k}\pi_h \leq r_k m, \quad \forall k \in \mathcal{K}, \tag{121}$$

$$n_{h,k} \geq 0, \quad \forall h \in \mathcal{H}, k\in\mathcal{K} \tag{122}$$

Note that the above linear program differs from FLP (Definition 6.1) only through changing $\ln(1/\delta)$ to $(1-\delta)\ln\frac{1-\delta}{\delta} - e^{-1}$ in Eq. (121). Hence, it is not difficult to verify that any value, m, associated with a feasible solution of FLP must satisfy

$$\frac{m}{m^*(\delta,\pi)} \geq \frac{(1-\delta)\ln\frac{1-\delta}{\delta} - e^{-1}}{\ln(1/\delta)} = 1-\delta - \frac{(1-\delta)\ln\frac{1}{1-\delta} + e^{-1}}{\ln(1/\delta)}, \tag{123}$$

for all sufficiently small δ, where $m^*(\delta,\pi)$ is the minimum system size. The completes the proof of Proposition 6.2. □

A.3 Proof of Lemma 6.3

Proof. The first inequality in Eq. (39) in fact holds for *any* feasible solution of FLP, and it follows directly from the constraints in Eq. (36) by noting that $D(h,l,k)$ is always no greater than $\overline{d}$. To show the second inequality in Eq. (39), fix some optimal solution $\{n^*_{h,k}\}$ such that $\frac{1}{\ln(1/\delta)}\sum_{k\in\mathcal{K}} n^*_{h,k} > c_{\mathcal{K}}\underline{d}^{-1}$, for some $h\in\mathcal{H}$. This would imply that there exists $k'\in\mathcal{K}$ such that $n^*_{h,k'} > \underline{d}^{-1}\ln(1/\delta)$. From the definition of $\underline{d}$ (Eq. (5)), we further conclude that, for all l such that $D(h,l,k')\neq 0$,

$$n^*_{h,k'} D(h,l,k') \geq n^*_{h,k'}\underline{d} > \ln(1/\delta). \tag{124}$$

Therefore, we may strictly reduce $n^*_{h,k'}$ without violating the second constraint (Eq. (36)) or increasing the objective value of FLP. This shows that we can decrease the value of $\frac{1}{\ln(1/\delta)} \sum_{k\in\mathcal{K}} n^*_{h,k}$ whenever it is strictly greater than $c_{\mathcal{K}}\underline{d}^{-1}$, and hence the second inequality in Eq. (39) must hold for some optimal solution of FLP. Finally, to show Eq. (40), we sum both sides of the the constraints in Eq. (35) over $\mathcal{K}$, and obtain

$$\begin{aligned}
m^*_F &= \sum_{k\in\mathcal{K}} r_k m^*_F \\
&\geq \sum_{k\in\mathcal{K}} \sum_{h\in\mathcal{H}} n^*_{h,k}\pi_h \\
&= \sum_{h\in\mathcal{H}} \pi_h \sum_{k\in\mathcal{K}} n^*_{h,k} \\
&\overset{(a)}{\geq} \overline{d}^{-1} \ln(1/\delta) \sum_{h\in\mathcal{H}} \pi_h \\
&= \overline{d}^{-1} \ln(1/\delta),
\end{aligned} \tag{125}$$

where step (a) follows from the first inequality in Eq. (39), which we have just shown. This proves Lemma 6.3. □

A.4 A Sufficient Condition for Accurate Classification

Fix $i, n \in \mathbb{N}$, and denote by $Y_{i,n} = \{(X_{i,j}, K_{i,j})\}_{j=1,\ldots,n}$, the inspection history of job i up till the nth inspection it receives, where the inspections are ranked according to their time of initiation. Let $S_{i,[n]}(h,l)$ be the cumulative log-likelihood from the first n inspections: $S_{i,[n]}(h,l) = \sum_{j=1}^{n} Z_{i,j}(h,l,K_{i,j})$, and similarly, let $\widehat{H}_{i,[n]}$ be the ML estimator of H_i from the first n inspections.

Lemma A.1 *Fix $i \in \mathbb{N}$ and $x > 0$. Let N^F_i be the total number of inspections job i receives before its departure from the system. Let $N \in \mathbb{N}$, $N \leq N^F_i$, be a stopping time with respect to Y_{i,N^F_i}. Denote by $\mathcal{G}_x$ the event:*

$$\mathcal{G}_x = \{\exists l \in \mathcal{H},\ s.t.\ S_{i,[N]}(l,h') \geq x,\ \forall h' \in \mathcal{H}, h' \neq l\}. \tag{126}$$

We have that

$$\mathbb{P}(\widehat{H}_{i,[N]} \neq h\,, \mathcal{G}_x \,\big|\, H_i = h) \leq c_{\mathcal{H}} \exp(-x), \quad \forall h \in \mathcal{H}. \tag{127}$$

Note that Lemma 4.1 follows from the above lemma by setting $N = N^F_i$.

Proof. Fix $i \in \mathbb{N}$. Let $y_n = \{(x_j, k_j)\}_{j=1,\ldots,n}$ be a particular realization of $Y_{i,n}$. We will use the notation of $q_h(y_n)$, defined in Eq. (102) in Appendix A.1, with a slight modification of changing N_i to N, i.e.,

$$q_h(y_n) = \mathbb{P}(N = n, Y_{i,n} = y_n \,\big|\, H_i = h). \tag{128}$$

Using a derivation that is identical to Eq. (104) in Appendix A.1, we have that

$$q_h(y_n) = v(y_n) \prod_{j=1}^{n} p(h, k_j, x_j). \tag{129}$$

where $v(y_n) \triangleq \mathbb{P}(N = n \,|\, Y_{i,n} = y_n) \prod_{j=1}^n \mathbb{P}(K_{i,j} = k_j, N \geq j \,|\, Y_{i,j-1} = y_{j-1})$, and it does not depend on h. We thus have that

$$\frac{q_h(Y_{i,n})}{q_l(Y_{i,n})} = \frac{\prod_{j=1}^n p(h, K_{i,j}, X_{i,j})}{\prod_{j=1}^n p(l, K_{i,j}, X_{i,j})} = \exp(-S_{i,[n]}(l,h)). \tag{130}$$

Fix $h, l \in \mathcal{H}$, $h \neq l$. Denote by $\mathcal{C}_n^l$ the event

$$\mathcal{C}_n^l = \{N = n\,,\, \widehat{H}_{i,[n]} = l\}. \tag{131}$$

In particular,

$$\mathbb{P}(\widehat{H}_{i,[N]} = l\,,\, \mathcal{G}_x \,|\, H_i = h) = \sum_{n=0}^{\infty} \mathbb{P}(C_n^l\,,\, \mathcal{G}_x \,|\, H_i = h). \tag{132}$$

We now employ a change-of-measure argument. Denote by $\mathcal{Y}_n$ the set of realizations, y_n, for which $\mathbb{P}(\mathcal{C}_n^l, Y_{i,n} = y_n \,|\, H_i = h) > 0$, and

$$S_{i,[n]}(l, h') \geq x, \quad \forall h' \in \mathcal{H}, h' \neq l. \tag{133}$$

We then obtain that

$$\begin{aligned}
\mathbb{P}(C_n^l\,,\, \mathcal{G}_x \,|\, H_i = h) &= \sum_{y_n \in \mathcal{Y}_n} q_h(y_n) \\
&= \sum_{y_n \in \mathcal{Y}_n} \frac{q_h(y_n)}{q_l(y_n)} q_l(y_n) \\
&\overset{(a)}{=} \sum_{y_n \in \mathcal{Y}_n} \exp(-S_{i,[n]}(l,h)) q_l(y_n) \\
&\overset{(b)}{\leq} \exp(-x) \sum_{y_n \in \mathcal{Y}_n} q_l(y_n) \\
&\leq \exp(-x) \mathbb{P}(C_n^l\,,\, \mathcal{G}_x \,|\, H_i = l),
\end{aligned} \tag{134}$$

where step (a) is based on Eq. (130), and step (b) from Eq. (133). Because Eq. (134) holds for any $l \neq h$ and $n \in \mathbb{N}$, using a union bound, we have that

$$\begin{aligned}
\mathbb{P}(\widehat{H}_i \neq h\,,\, \mathcal{G}_x \,|\, H_i = h) &= \sum_{l \in \mathcal{H}, l \neq h} \mathbb{P}(\widehat{H}_i = l\,,\, \mathcal{G}_x \,|\, H_i = h) \\
&\overset{(a)}{=} \sum_{l \in \mathcal{H}, l \neq h} \left(\sum_{n=0}^{\infty} \mathbb{P}(C_n^l\,,\, \mathcal{G}_x \,|\, H_i = h) \right) \\
&\overset{(b)}{\leq} \sum_{l \in \mathcal{H}, l \neq h} \exp(-x) \left(\sum_{n=0}^{\infty} \mathbb{P}(C_n^l\,,\, \mathcal{G}_x \,|\, H_i = l) \right) \\
&= \exp(-x) \sum_{l \in \mathcal{H}, l \neq h} \mathbb{P}(\widehat{H}_{i,[n]} = l \,|\, H_i = l) \\
&\leq c_{\mathcal{H}} \exp(-x),
\end{aligned} \tag{135}$$

where step (a) follows from Eq. (132), and step (b) from Eq. (134). This proves our claim. □

A.5 Proof of Proposition 7.2

Proof. Fix $h \in \mathcal{H}$ and $i \in \mathbb{N}$. Denote by $S_i^P(h,l)$ the value of $S_{i,t}(h,l)$ at the time when job i exits the Preparation stage. Define the event

$$\mathcal{B} = \left\{ S_i^P(h,l) \geq \ln\ln(1/\delta),\ \forall l \in \mathcal{H}\backslash\{h\} \right\}. \tag{136}$$

The following lemma is the main technical result that we will use. Its proof makes use of the Azuma-Hoeffding inequality by noting that $S_{i,t}(h,l)$ is a martingale under proper conditioning.

Lemma A.2 *Denote by $\overline{\mathcal{B}}$ the complement of $\mathcal{B}$. For all $h \in \mathcal{H}$,*

$$\mathbb{P}\left(\overline{\mathcal{B}} \,\middle|\, H_i = h\right) \leq c_{\mathcal{H}} \ln^{-1}(1/\delta). \tag{137}$$

Proof. We will index all inspections received by job i in the Preparation stage in an arbitrary fashion, and denote by K_j the type of the expert who performed the jth inspection for job i, and by $Z_j(\cdot,\cdot,K_j)$ the corresponding log-likelihood ratio. Define

$$M_n^l = \sum_{s=1}^{n \vee n^P} \left(Z_s(h,l,K_s) - d(h,l) \right), \quad n \in \mathbb{N}, \tag{138}$$

where

$$d(h,l) = \sum_{k \in \mathcal{K}} D(h,l,k) r_k. \tag{139}$$

Recall that the experts visit the three stages according to the randomized rule described in Section 5.1, and hence the probability of the jth expert to inspect job i being type k is equal to r_k. Conditional on the true label of job i being h, we have that the $Z_s(h,l,K_s)$'s are i.i.d., with $\mathbb{E}(Z_1(h,l,K_1)|H_i = h) = d(h,l)$. It is hence not difficult to verify that $\{M_n^l \,|\, H_i = h\}_{n\in\mathbb{N}}$ is a martingale. Recall that $d_{\mathrm{a}} \triangleq \min_{h,l\in\mathcal{H},h\neq l} d(h,l)$. We have that for any $l \in \mathcal{H}\backslash\{h\}$,

$$\begin{aligned}
\mathbb{P}(S_i^P(h,l) \leq \ln\ln(1/\delta) \,|\, H_i = h) =& \mathbb{P}(S_i^P(h,l) - n^P d(h,l) \leq \ln\ln(1/\delta) - n^P d(h,l) \,|\, H_i = h) \\
\leq& \mathbb{P}(S_i^P(h,l) - n^P d(h,l) \leq \ln\ln(1/\delta) - n^P d_{\mathrm{a}} \,|\, H_i = h) \\
=& \mathbb{P}(S_i^P(h,l) - n^P d(h,l) \leq -(d_{\mathrm{a}} - \ln\ln(1/\delta)/n^P) n^P \,|\, H_i = h) \\
\overset{(a)}{\leq}& \exp\left(-\frac{(d_{\mathrm{a}} - \ln\ln(1/\delta)/n^P)^2 (n^P)^2}{8\overline{z}^2 n^P} \right) \\
\leq& \exp\left(-\frac{d_{\mathrm{a}}^2 n^P - 2 d_{\mathrm{a}} \ln\ln(1/\delta)}{8\overline{z}^2} \right) \\
\overset{(b)}{=}& \exp\left(-\ln\ln(1/\delta)\right) \\
=& \ln^{-1}(1/\delta).
\end{aligned} \tag{140}$$

Step (a) follows from the Azuma-Hoeffding's Inequality (Lemma C.1 in Appendix C), and that $|M_n^l - M_{n-1}^l| \leq 2\overline{z}$ for all $n \in \mathbb{N}$. Step (b) follows from the fact that

$$n^P = \zeta_0 \ln\ln(1/\delta) = \frac{8\overline{z}^2 + 2d_{\mathrm{a}}}{d_{\mathrm{a}}^2} \ln\ln(1/\delta). \tag{141}$$

Because Eq. (140) holds for all l, Lemma A.2 follows by applying a union bound over $l \in \mathcal{H}$. $\square$

We now return to the proof of Proposition 7.2. Combining Lemmas 4.1 and A.2, we have that

$$\begin{aligned}\mathbb{P}(\widehat{H}_i^P \neq h \,|\, H_i = h) \leq& \mathbb{P}(\overline{\mathcal{B}} \,|\, H_i = h) + \mathbb{P}(\widehat{H}_i^P \neq h \,,\, \mathcal{B} \,|\, H_i = h)\\ \stackrel{(a)}{\leq}& c_{\mathcal{H}} \ln^{-1}(1/\delta) + \mathbb{P}(\widehat{H}_i^P \neq h \,,\, \mathcal{B} \,|\, H_i = h)\\ \stackrel{(b)}{\leq}& c_{\mathcal{H}} \ln^{-1}(1/\delta) + c_{\mathcal{H}} \exp(-\ln\ln(1/\delta))\\ =& 2c_{\mathcal{H}} \ln^{-1}(1/\delta),\end{aligned} \tag{142}$$

where step (a) follows from Lemma A.2, and (b) from Lemma A.1 by setting $x = \ln\ln(1/\delta)$, $N = n^P$, and noting that $\mathcal{B} \subset \mathcal{G}_x$. Our claim that $\epsilon^P \leq 2c_{\mathcal{H}} \ln^{-1}(1/\delta)$ follows by noting that the above equation holds for all $h \in \mathcal{H}$. To show Eq. (43), note that, for all $h \in \mathcal{H}$,

$$\begin{aligned}\pi_h^P \leq& \pi_h + \sum_{h' \neq h} \pi_{h'} \mathbb{P}(\widehat{H}_1^P = h \,|\, H_1 = h')\\ \leq& \pi_h + \epsilon^P \sum_{h' \neq h} \pi_{h'}\\ \leq& \pi_h + 2c_{\mathcal{H}} \ln^{-1}(1/\delta).\end{aligned} \tag{143}$$

where the last inequality follows from the upper bound on ϵ^P, and that $\sum_{h' \neq h} \pi_{h'} \leq 1$. This completes the proof of Proposition 7.2. $\square$

A.6 Proof of Proposition 7.7

Proof. We start by looking at the workload process for the Preparation stage, $\mathbf{W}_0^s(t)$. For every $a, b > 0$, with $a > b$, $s \in \mathbb{N}$, and $z > 0$, applying the triangle inequality to Eq. (54), we have that

$$\begin{aligned}|\mathbf{W}_0^s(z,a) - \mathbf{W}_0^s(z,b)| \leq& |\mathbf{A}_0^s(z,a) - \mathbf{A}_0^s(z,b)| + |\mathbf{\Delta}_0^s(z,a) - \mathbf{\Delta}_0^s(z,b)|\\ \stackrel{(a)}{=}& n^P |\Xi_0(z,a) - \Xi_0(z,b)| + |\mathbf{\Delta}_0^s(z,a) - \mathbf{\Delta}_0^s(z,b)|\\ \stackrel{(b)}{\leq}& n^P |\Xi_0(z,a) - \Xi_0(z,b)| + \sum_{e \in \mathcal{E}} |R_{e,P}(z,a) - R_{e,P}(z,b)|.\end{aligned} \tag{144}$$

where $\Xi_0(\cdot)$ and $R_{e,T}(\cdot)$ are defined in Eqs. (51) and (52), respectively. Step (a) follows from the design of our inspection policy, where each job that arrives to the Preparation stage adds n^P new inspections to $\mathbf{W}_0(t)$. Step (b) is due to the fact that the number of initiated inspections during an interval is no greater than the total number of times that an expert visits the Preparation stage during the same period.

Recall that the average inspection rate, $\overline{\mu}$, is equal to 1 (Eq. (2)). Because $\omega \in \mathcal{C}$ and $\lim_{s \to \infty} z_s = \infty$, by Eqs. (51), (52) and (144), we have that there exists a positive sequence $\{\epsilon_s\}_{s \in \mathbb{N}}$, with $\epsilon_s \to 0$ as $s \to \infty$, such that for all $s \in \mathbb{N}$ and $a, b \in [0, T], a > b$,

$$|\mathbf{A}_0^s(z_s,a) - \mathbf{A}_0^s(z_s,b)| \leq n^p(a-b) + \epsilon_s, \tag{145}$$

$$|\mathbf{\Delta}_0^s(z_s,a) - \mathbf{\Delta}_0^s(z_s,b)| \leq q^P m(a-b) + \epsilon_s, \tag{146}$$

$$|\mathbf{W}_0^s(z_s,a) - \mathbf{W}_0^s(z_s,a)| \leq (n^P + q^P m)(a-b) + \epsilon_s. \tag{147}$$

Recall that $\mathbf{W}_0(0) = \mathbf{w}^{(s)}$ in the sth system, and $\lim_{s\to\infty} {z_s}^{-1}\mathbf{w}^{(s)} = \mathbf{w}^0$. We invoke the following variation of the Arzelà-Ascoli theorem (Lemma 6.3 of Ye et al. (2005)).

Lemma A.3 *Fix $T > 0$. Let $\{f_s(\cdot)\}_{s\in\mathbb{N}}$ a sequence of functions that satisfies*

1. $\sup_{s\in\mathbb{N}} |f_s(0)| < \infty$.
2. *There exists a positive sequence $\{\epsilon_s\}_{s\in\mathbb{N}}$ with $\epsilon_s \to 0$ as $s \to \infty$, such that*

$$|f_s(a) - f_s(b)| \le c|a-b| + \epsilon_s, \quad \forall a, b \in [0,T],\ s \in \mathbb{N}. \tag{148}$$

Then, there exists a c-Lipschitz function $f^*(\cdot)$, and an increasing sequence $\{s_j\}_{j\in\mathbb{N}} \subset \mathbb{N}$, such that

$$\lim_{j\to\infty} \left\| f_{s_j}(t) - f^*(t) \right\|_T = 0. \tag{149}$$

Using Lemma A.3 and Eqs. (145) through (147), we have that there exist $(n^P + q^P m)$-Lipschitz functions, $\mathbf{a}_0(\cdot)$, $\mathbf{d}_0(\cdot)$, and $\mathbf{w}_0(\cdot)$, and an increasing sequence $\{i_j^0\}_{j\in\mathbb{N}}$, such that

$$\lim_{j\to\infty} \left\| (\mathbf{W}_0, \mathbf{A}_0, \mathbf{\Delta}_0)^{s_{i_j^0}}(z_{s_{i_j^0}}, t) - (\mathbf{w}_0, \mathbf{a}_0, \mathbf{d}_0)(t) \right\|_T = 0. \tag{150}$$

This shows that the sequence $\{(\mathbf{W}_0, \mathbf{A}_0, \mathbf{\Delta}_0)^{s_i}(z_{s_i}, \cdot)\}_{i\in\mathbb{N}}$ admits Lipschitz-continuous limit points.

We now repeat essentially the same argument for the workload in the Adaptive stage, and use a diagonal argument to identify a family of nested sub-sequences along which the workload converges coordinate-wise to a Lipschitz function. Fix $k \in \{1, \ldots, \mathcal{K}\}$. We claim that

$$|\mathbf{A}_k^s(a) - \mathbf{A}_k^s(b)| \le v_\delta \left(m + \sum_{e\in\mathcal{E}} |R_{e,P}(a) - R_{e,P}(b)| \right), \quad \forall a, b \in [0,T], a > b. \tag{151}$$

To justify the above inequality, note that by Eq. (28), each job that arrives to the Adaptive stage could add at most v_δ inspections to the aggregate workload in the Adpative stage. Also, an arrival to the Adaptive stage corresponds to a departure to the Preparation stage, and the number of such departures during an interval is at most the number of times an expert visits the Preparation stage, $\sum_{e\in\mathcal{E}} |R_{e,P}(a) - R_{e,P}(b)|$, plus the number of experts who were already inspecting a job in the Preparation stage prior to time $t = a$, which is at most m. This yields Eq. (151). Analogous to Eq. (144), we have that

$$\begin{aligned} |\mathbf{W}_k^s(z,a) - \mathbf{W}_0^s(z,b)| \le & n^P |\mathbf{A}_k^s(z,a) - \mathbf{A}_k^s(z,b)| + |\mathbf{\Delta}_k^s(z,a) - \mathbf{\Delta}_k^s(z,b)| \\ \le & v_\delta \left(\frac{m}{z} + \sum_{e\in\mathcal{E}} |R_{e,P}(z,a) - R_{e,P}(z,b)| \right) + \sum_{e\in\mathcal{E}} |R_{e,A}(z,a) - R_{e,A}(z,b)|. \end{aligned} \tag{152}$$

Using the fact that $\omega \in \mathcal{C}$, $\lim_{s\to\infty} z_s = \infty$, and Eqs. (52) and (152), we have that there exists a positive sequence $\{\epsilon_s\}_{s\in\mathbb{N}}$, with $\lim_{s\to\infty} \epsilon_s = 0$, such that, for all $s \in \mathbb{N}$ and $a, b \in [0,T], a > b$,

$$|\mathbf{A}_k^s(z_s,a) - \mathbf{A}_k^s(z_s,b)| \le v_\delta q^P m(a-b) + \epsilon_s, \tag{153}$$

$$|\mathbf{\Delta}_k^s(z_s,a) - \mathbf{\Delta}_k^s(z_s,b)| \le q^A m(a-b) + \epsilon_s, \tag{154}$$

$$|\mathbf{W}_k^s(z_s,a) - \mathbf{W}_k^s(z_s,b)| \le (v_\delta q^P + q^A) m(a-b) + \epsilon_s. \tag{155}$$

Using Eq. (155) and the same argument involving Lemma A.3 as before, we have that there exist

1. a collection of coordinate-wise $(v_\delta q^P + q^A)m$-Lipschitz functions, $\{(\mathbf{w}_k, \mathbf{a}_k, \mathbf{d}_k)\}_{k\in\mathcal{K}}$, and
2. a family of sequences, $\{\{i_j^k\}_{j\in\mathbb{N}}\}_{k\in\mathcal{K}}$, where $\{i_j^k\}$ is a sub-sequence of $\{i_j^{k-1}\}$ for all $k \in \mathcal{K}$, and $\{i_j^0\}$ was defined in Eq. (150),

such that

$$\lim_{j\to\infty} \left\| (\mathbf{W}_k, \mathbf{A}_k, \mathbf{\Delta}_k)^{s_{i_j^k}} (z_{s_{i_j^k}}, t) - (\mathbf{w}_k, \mathbf{a}_k, \mathbf{d}_k)(t) \right\|_T = 0, \quad \forall k \in \{1, \ldots, c_{\mathcal{K}}\}. \tag{156}$$

Because the sequences, $\{i_j^k\}_{k\in\mathcal{K}}$, are nested, by setting $k = c_{\mathcal{K}}$, the above equation further implies that

$$\lim_{j\to\infty} \left\| (\mathbf{W}, \mathbf{A}, \mathbf{\Delta})^{s_{i_j^{c_{\mathcal{K}}}}} (z_{s_{i_j^k}}, t) - (\mathbf{w}, \mathbf{a}, \mathbf{d})(t) \right\|_T = 0, \tag{157}$$

Note that, by construction, all coordinates of $(\mathbf{w}, \mathbf{a}, \mathbf{d})$ are Lipschitz-continuous, with a Lipschitz constant of $c = [q^A + (v_\delta + 1)q^P]m + n^P$. We complete the proof of Proposition 7.7 by setting $\{i_j\}_{j\in\mathbb{N}} = \{i_j^{c_{\mathcal{K}}}\}_{j\in\mathbb{N}}$. □

A.7 Proof of Lemma 7.9

Proof. The general strategy of our proof is to relate the behavior of $B_h^s(\cdot)$ to the number of initiated inspections in the Preparation stage, $\Delta_0^s(\cdot)$, while being cautious of the fact that the lengths of inspections are random and can cause the order of job departures from the Preparation stage to be different from that of the arrivals. Let $B^s(t) = |\mathcal{B}^s(t)|$. Fix $t \in (0, T)$ and $t' \in (t, T)$. We first show that the following property holds:

$$\left| (B^s(t') - B^s(t)) - (\Delta_0^s(t') - \Delta_0^s(t))/n^P \right| \le 2m. \tag{158}$$

In words, this means that the number of jobs that enter the Adaptive stage during the interval $(t, t']$ multiplied by n^P deviates from the number of initiated inspections during the same period by at most $2mn^P$.

To prove Eq. (158), recall that each job must receive n^P inspections in the Preparation stage before it moves to the Adaptive stage, and hence having $(\Delta_0^s(t') - \Delta_0^s(t))$ initiated inspections means that there could be at most $\lceil (\Delta_0^s(t') - \Delta_0^s(t))/n^P \rceil + m$ jobs arriving at the Adaptive stage during $(t, t']$, where the addition of m captures the possibility that all m servers were processing a job from the Preparation stage at the start of the interval, t. This shows that

$$(B^s(t') - B^s(t)) - (\Delta_0^s(t') - \Delta_0^s(t))/n^P \le m + 1 \le 2m. \tag{159}$$

We next invoke the following fact:

Lemma A.4 *Fix $x \in \mathbb{N}$ and an interval $J \subset \mathbb{R}_+$. If there are x inspections initiated in the Preparation stage during J, then there are at least $(x/n^P - m)$ jobs departing from the Preparation stage during J.*

Proof. Denote by $\mathcal{I}'$ the set of jobs who have had *any* inspection initiated during the interval J. We partition $\mathcal{I}'$ into $\mathcal{I}'_0 \cup \mathcal{I}'_1$, where $\mathcal{I}'_0$ corresponds to those jobs who have departed from the Preparation stage by the end of J, and $\mathcal{I}'_1$ to those who have not. Because the experts perform inspections in a first-come-first-server manner, a job cannot start receiving inspections until all n^P inspections for the previous job have been initiated. Since there are m experts, this implies that at any point in time, there can be at most m jobs in the Preparation stage who have initiated any inspection, and therefore

$$|\mathcal{I}'_1| \leq m. \tag{160}$$

Recall that x is the total number of initiated inspections in the Preparation stage during J. We have that

$$x \overset{(a)}{\leq} n^P|\mathcal{I}'| = n^P(|\mathcal{I}'_0| + |\mathcal{I}'_1|) \overset{(b)}{\leq} n^P(|\mathcal{I}'_0| + m), \tag{161}$$

where step (a) follows from the fact that each job can lead to at most n^P inspections, and step (b) from Eq. (160). This yields

$$|\mathcal{I}'_0| \geq x/n^P - m. \tag{162}$$

That is, there are at least $(x/n^P - m)$ jobs departing from the Preparation stage during J. $\square$

We now apply Lemma A.4 with $x = (\Delta_0^s(t') - \Delta_0^s(t))$, and obtain

$$B^s(t') - B^s(t) \geq (\Delta_0^s(t') - \Delta_0^s(t))/n^P - m, \tag{163}$$

or

$$(B^s(t') - B^s(t)) - (\Delta_0^s(t') - \Delta_0^s(t))/n^P \geq -m, \tag{164}$$

which, combined with Eq. (159) leads to Eq. (158).

To prove the main claim of Lemma 7.9, we will invoke the functional law of large number on the sequence of ML estimators, $\{\widehat{H}_i^P\}_{i\in\mathbb{N}}$, as expressed in Eq. (53). However, we have to be careful in doing so: the jobs that depart from the Preparation stage do not necessarily preserve the order in which they arrived to the system, since the lengths of inspections in the Preparation stage are random and can potentially cause a job to depart earlier than another with a smaller index. Thus, we begin by showing that such "shuffling" is relatively minor and can essentially be ignored when studying the local behavior of $B_h^s(\cdot)$. This is formalized in the next lemma, which states that the set $\mathcal{B}^s(t')\backslash\mathcal{B}^s(t)$ can be "sandwiched" by two intervals of consecutive integers that differ by no more than $2m$ elements from one another. The use of consecutive integers will allow us to invoke the law of large numbers with greater ease.

We will adopt the following short-hand notation. For a set-valued process $\{\mathcal{S}(t)\}_{t\in\mathbb{R}_+}$, and $t, t' \in \mathbb{R}_+$, we will denote by $\mathcal{S}(t'\backslash t)$ the difference between $\mathcal{S}(t')$ and $\mathcal{S}(t)$:

$$\mathcal{S}(t'\backslash t) \overset{\triangle}{=} \mathcal{S}(t')\backslash\mathcal{S}(t). \tag{165}$$

For $a, b \in \mathbb{N}$, we will use $\{a \to b\}$ to denote the set of consecutive integers $\{a, a+1, \ldots, b\}$, if $a \leq b$, and $\emptyset$, otherwise. We have the following lemma.

Lemma A.5 *Fix* $t, t' \in \mathbb{R}_+$, $t' > t$. *There exist* a, a', b *and* $b' \in \mathbb{N}$, *with* $\max\{|a-a'|, |b-b'|\} \leq m$, *such that*

$$\{a \to b\} \subset \mathcal{B}^s(t' \backslash t) \subset \{a' \to b'\}. \tag{166}$$

It follows that

$$\left|(b-a) - |\mathcal{B}^s(t' \backslash t)|\right| \leq 2m, \text{ and } \left|(b'-a') - |\mathcal{B}^s(t' \backslash t)|\right| \leq 2m. \tag{167}$$

Proof. Denote by $\mathcal{D}^s(t)$ the set of jobs in the Preparation stage for whom at least one inspection has been initiated by time t. Because a job can depart from the Preparation stage only after completing n^P inspections, we have

$$\mathcal{B}^s(t) \subset \mathcal{D}^s(t), \quad \forall t \in \mathbb{R}_+. \tag{168}$$

On the other hand, the initiations of inspections in the Preparation stage are performed in a first-come-first-serve (FCFS) manner. Recall also that at any given point in time there can be at most m experts performing inspections. Combining the above two facts, we know that an inspection for job $i+m$ can be initiated only after all jobs prior to, and including, job i have departed from the Preparation stage. We thus have that

$$\{1 \to \max \mathcal{D}^s(t) - m\} \subset \mathcal{B}^s(t), \quad \forall t \in \mathbb{R}_+. \tag{169}$$

where $\max \mathcal{D}^s(t)$ is the index of the last job in $\mathcal{D}^s(t)$. Combining Eqs. (168) and (169), we have

$$\mathcal{B}^s(t' \backslash t) \supset \{1 \to \max \mathcal{D}^s(t') - m\} \backslash \mathcal{D}^s(t), \tag{170}$$
$$\mathcal{B}^s(t' \backslash t) \subset \mathcal{D}^s(t') \backslash \{1 \to \max \mathcal{D}^s(t) - m\} \tag{171}$$

Finally, the above-mentioned FCFS property of the inspection initiation rule further implies that $\mathcal{D}^s(t)$ is a set of consecutive positive integers. Therefore, it follows that the right-hand sides of Eqs. (170) and (171) are both intervals of consecutive integers. We can prove Eq. (166) by setting

$$\{a \to b\} = \{1 \to \max \mathcal{D}^s(t') - m\} \setminus \mathcal{D}^s(t), \tag{172}$$
$$\{a' \to b'\} = \mathcal{D}^s(t') \setminus \{1 \to \max \mathcal{D}^s(t) - m\}, \tag{173}$$

or, equivalently, that

$$\begin{aligned} a &= \max \mathcal{D}^s(t) + 1, \quad a' = \max \mathcal{D}^s(t) + 1 - m \\ b &= \max \mathcal{D}^s(t') - m + 1, \quad b' = \max \mathcal{D}^s(t') + 1 \end{aligned} \tag{174}$$

and it is clear from the above equations that $\max\{|a-a'|, |b-b'|\} \leq m$. For Eq. (167), note that by the ordering in Eq. (166) and the triangle inequality, we have

$$\left|(b-a) - |\mathcal{B}^s(t' \backslash t)|\right| \leq \left|(b'-a') - (b-a)\right| \leq 2\max\{|a-a'|, |b-b'|\} \leq 2m, \tag{175}$$

and a similar argument shows $\left|(b'-a') - |\mathcal{B}^s(t' \backslash t)|\right| \leq 2m$. This proves Lemma A.5. □

We are now ready to prove the main claim of Lemma 7.9. Fix $t \in (0, T)$ and $\epsilon > 0$. Let a, a', b and b' be defined as in Lemma A.5, corresponding to the set $\mathcal{B}^{\overline{s}_j}(\overline{z}_j(t+\epsilon) \backslash \overline{z}_j t)$, all of which depend on j and ϵ, though we will suppress the dependencies in our notation for simplicity. We have that, for every sufficiently small $\delta' > 0$, there exists $j^* > 0$, such that for all $j > j^*$,

$$
\begin{aligned}
\frac{B_h^{\bar{s}_j}(\bar{z}_j, t+\epsilon) - B_h^{\bar{s}_j}(z_{\bar{s}_j}, t)}{\epsilon} &= \frac{B_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon)) - B_h^{\bar{s}_j}(z_{\bar{s}_j} t)}{\epsilon \bar{z}_j} \\
&= \frac{1}{\epsilon \bar{z}_j} \sum_{i \in \mathcal{B}^{\bar{s}_j}(\bar{z}_j(t+\epsilon)\setminus \bar{z}_j t)} \mathbb{I}(\widehat{H}_i^P = h) \overset{(a)}{\geq} \frac{1}{\epsilon \bar{z}_j} \sum_{i=a}^{b} \mathbb{I}(\widehat{H}_i^P = h) \\
&\overset{(b)}{\geq} \frac{1}{\epsilon \bar{z}_j} \sum_{i=a}^{a+(B^{\bar{s}_j}(\bar{z}_j(t+\epsilon) - B^{\bar{s}_j}(\bar{z}_j t)) - 2m} \mathbb{I}(\widehat{H}_i^P = h) \\
&\overset{(c)}{\geq} \frac{1}{\epsilon \bar{z}_j} \sum_{i=a}^{a+(\boldsymbol{\Delta}_0^{\bar{s}_j}(\bar{z}_j(t+\epsilon) - \boldsymbol{\Delta}_0^{\bar{s}_j}(\bar{z}_j t))/n^P - 4m} \mathbb{I}(\widehat{H}_i^P = h) \\
&\overset{(d)}{\geq} \frac{1}{\epsilon \bar{z}_j} \sum_{i=a}^{a+(\mathbf{d}_0(t+\epsilon) - \mathbf{d}_0(t) - \delta')\bar{z}_j/n^P - 4m} \mathbb{I}(\widehat{H}_i^P = h) \\
&\overset{(e)}{\geq} \frac{\pi_h^P}{\epsilon n^P} (\mathbf{d}_0(t+\epsilon) - \mathbf{d}_0(t) - 2\delta') \\
&= \frac{\pi_h^P}{n^P} \cdot \frac{\mathbf{d}_0(t+\epsilon) - \mathbf{d}_0(t) - 2\delta'}{\epsilon}. \qquad (176)
\end{aligned}
$$

Step (a) follows from Eq. (166) of Lemma A.5, and step (b) from the first inequality in Eq. (167) in the same lemma. Step (c) is a consequence of $\left|(B^s(t') - B^s(t)) - (\boldsymbol{\Delta}_0^s(t') - \boldsymbol{\Delta}_0^s(t))/n^P\right| \leq 2m$ (Eq. (158)). Step (d) follows from the uniform convergence of $\boldsymbol{\Delta}_0^{\bar{s}_j}(\cdot)$ to $\mathbf{d}_0(\cdot)$ over $[0, T]$. Finally, step (e) is based on the functional law of large number for the sequence $\{\widehat{H}_i^P\}_{i \in \mathbb{N}}$ (Eq. (53)).

Since Eq. (176) holds for all sufficiently small $\delta' > 0$, we have that

$$
\begin{aligned}
\lim_{\epsilon \downarrow 0} \lim_{j \to \infty} \frac{B_h^{\bar{s}_j}(\bar{z}_j, t+\epsilon) - B_h^{\bar{s}_j}(z_{\bar{s}_j}, t)}{\epsilon} &\geq \lim_{\epsilon \downarrow 0} \lim_{j \to \infty} \lim_{\delta' \downarrow 0} \frac{\pi_h^P}{n^P} \cdot \frac{\mathbf{d}_0(t+\epsilon) - \mathbf{d}_0(t) - 2\delta'}{\epsilon} \\
&= \frac{\pi_h^P}{n^P} \dot{\mathbf{d}}_0(t) \\
&= \begin{cases} \frac{mq^P}{n^P} \pi_h^P, & \text{if } \mathbf{w}_0(t) > 0, \\ \pi_h^P, & \text{if } \mathbf{w}_0(t) = 0, \end{cases} \qquad (177)
\end{aligned}
$$

A line of argument identical to Eqs. (176) through (177) shows the other direction of the inequality, namely, that $\lim_{\epsilon \downarrow 0} \lim_{j \to \infty} \frac{B_h^{\bar{s}_j}(\bar{z}_j, t+\epsilon) - B_h^{\bar{s}_j}(z_{\bar{s}_j}, t)}{\epsilon} \leq \frac{\pi_h^P}{n^P} \dot{\mathbf{d}}_0(t)$. This completes the proof of Lemma 7.9. $\square$

A.8 Proof of Lemma 7.10

Proof. We first state a useful technical lemma. The proof makes use of Berge's maximum theorem concerning the continuity of optimal solutions to a convex optimization problem.

Lemma A.6 *Fix $h \in \mathcal{H}$ and $\mathbf{w} \in \mathbb{R}_+^{c\mathcal{K}+1}$. The set-valued function, $\mathcal{N}_h^*(\cdot)$, satisfies the following semi-continuity property. For every $\delta > 0$, there exists $\epsilon > 0$ such that, for all $\mathbf{w}' \in \mathbb{R}_+^{c\mathcal{K}+1}$, $\|\mathbf{w}' - \mathbf{w}\|_2 \le \epsilon$,*

$$\sup_{x \in \mathcal{N}_h^*(\mathbf{w}')} \inf_{y \in \mathcal{N}_h^*(\mathbf{w})} \|x - y\|_2 \le \delta. \tag{178}$$

Proof. The set $\mathcal{N}_h(\cdot)$ represents the set of feasible solutions of the linear optimization problem defined in Eq. (25). We first invoke Berge's maximum theorem (pp. 116 of Berge (1963)), which roughly states that $\mathcal{N}_h(\cdot)$ depends semi-continuously in its argument. More specifically, it is easy to verify that the set, $\mathcal{N}_h$, is compact (Eq. (26) to (28)), and the objective function is continuous. Berge's maximum theorem thus implies that $\mathcal{N}_h^*(\cdot)$ is upper-hemicontinuous at $\mathbf{w}$, in the following sense: let $\{\mathbf{w}_n\}_{n\in\mathbb{N}}$ be a sequence, where $\mathbf{w}_n \to \mathbf{w}$ as $n \to \infty$, and $\{\lambda_n\}_{n\in\mathbb{N}}$ be a sequence where $\lambda_n \in \mathcal{N}_h^*(\mathbf{w}^n)$ for all $n \in \mathbb{N}$. Then, if $\lambda_n \to \lambda$ as $n \to \infty$, we have that $\lambda \in \mathcal{N}_h^*(\mathbf{w})$.

We now use the above upper hemicontinuity property, along with the boundedness of the constraint set, $\mathcal{N}_h$, to prove our claim. Suppose, for the sake of contradiction, that there exists a sequence $\{\mathbf{w}_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \mathbf{w}_n = \mathbf{w}$, such that

$$\liminf_{n\to\infty} \sup_{x \in \mathcal{N}_h^*(\mathbf{w}_n)} \inf_{y \in \mathcal{N}_h^*(\mathbf{w})} \|x - y\|_2 > 0. \tag{179}$$

This implies the existence of a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$, with $\lambda_n \in \mathcal{N}_h^*(\mathbf{w}^n)$ for all $n \in \mathbb{N}$, such that

$$\liminf_{n\to\infty} \inf_{y \in \mathcal{N}_h^*(\mathbf{w})} \|\lambda_n - y\|_2 > 0. \tag{180}$$

By definition, for every $\mathbf{w}$, the coordinates of the elements of $\mathcal{N}_h^*(\mathbf{w})$ are non-negative and bounded from above by v_δ. Therefore, by sequential compactness, there exist $\lambda \in \mathbb{R}_+^{c\mathcal{K}+1}$ and a sub-sequence of $\{\lambda_n\}$, $\{\lambda_{n_k}\}_{k\in\mathbb{N}}$, such that $\lim_{k\to\infty} \lambda_{n_k} = \lambda$. By Eq. (180), we have that

$$\inf_{y \in \mathcal{N}_h^*(\mathbf{w})} \|\lambda - y\|_2 > 0. \tag{181}$$

This contradicts with the hemicontinuity of $\mathcal{N}_h^*(\cdot)$, which would imply that $\lambda \in \mathcal{N}_h^*(\mathbf{w})$, and thus proves Lemma A.6. ☐

We now return to the proof of Lemma 7.10. Fix $\epsilon \in (0, T - t)$. We first observe that the set $\mathcal{N}_h^*(\cdot)$ is scale-invariant, in the sense that for every $\mathbf{w} \in \mathbb{R}_+^{c\mathcal{K}+1}$

$$\mathcal{N}_h^*(a\mathbf{w}) = \mathcal{N}_h^*(\mathbf{w}), \quad a > 0. \tag{182}$$

Therefore,

$$\mathcal{N}_h^*(\mathbf{W}^s(zt)) = \mathcal{N}_h^*(z^{-1}\mathbf{W}^s(zt)) = \mathcal{N}_h^*(\mathbf{W}^s(z,t)), \tag{183}$$

Recall that $\mathbf{W}^{s_j}(\bar{z}_j, \cdot)$ converges uniformly over $[0, T]$ to $\mathbf{w}(\cdot)$ as $j \to \infty$, and that $\mathbf{w}(\cdot)$ is a continuous function. These two facts, along with the semi-continuity property of Lemma A.6, imply that for every $\delta > 0$, there exist $\epsilon, j^* > 0$, such that for all $j \ge j^*$

$$\inf_{y \in \mathcal{N}_h^*(\mathbf{w}(t))} \|\lambda - y\|_2 \le \delta, \quad \forall t' \in (t, t+\epsilon],\, \lambda \in \mathcal{N}_h^*(\mathbf{W}^{\bar{s}_j}(\bar{z}_j, t')). \tag{184}$$

Note that all jobs in the set $\mathcal{B}_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon)\backslash \bar{z}_j t)$ arrived during the interval $(\bar{z}_j t, \bar{z}_j(t+\epsilon)]$. Therefore, for all $i \in \mathcal{B}_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon)\backslash \bar{z}_j t)$, there exists $t' \in (t, t+\epsilon]$, such that

$$\Lambda_i \le \mathcal{N}_h^*(\mathbf{W}^{\bar{s}_j}(\bar{z}_j, t')), \tag{185}$$

where the inequality follows from round-down procedure in Eq. (31).

Combining Eqs. (184) and (185), we conclude that for every $\delta > 0$, there exist $\epsilon^*, j^* > 0$, such that for all $j \geq j^*, \epsilon < \epsilon^*$,

$$\inf_{y \in \mathcal{N}_h^*(\mathbf{w}(t))} \max_{k \in \mathcal{K}} (\Lambda_{i,k} - y_k) \leq \delta, \quad \forall i \in \mathcal{B}_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon) \backslash \bar{z}_j t). \tag{186}$$

Since the above inequality holds for *all* jobs in $\mathcal{B}_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon)\backslash \bar{z}_j t)$, it further implies that the average workload among the jobs in $\mathcal{B}_h^{\bar{s}_j}(\bar{z}_j(t+\epsilon)\backslash \bar{z}_j t)$ satisfies:

$$\inf_{y \in \mathcal{N}_h^*(\mathbf{w}(t))} \max_{k \in \mathcal{K}} \left(\overline{\Lambda}_{h,k}(t, \epsilon, j) - y_k\right) \leq \delta. \tag{187}$$

Because the above inequalities hold for all $\delta > 0$, we conclude that

$$\limsup_{\epsilon \downarrow 0} \limsup_{j \to \infty} \inf_{y \leq \mathcal{N}_h^*(\mathbf{w}(t))} \|\overline{\Lambda}_h(t, \epsilon, j) - y\|_2 = 0. \tag{188}$$

This completes the proof of Lemma 7.10. □

A.9 Proof of Proposition 7.11

Proof. Fix $\mathbf{w}^0 \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$ such that $L(\mathbf{w}^0) = 1$, and a fluid solution $\mathbf{w} \in \mathcal{W}(\mathbf{w}^0)$. Fix $t \in (0, T)$ to be a point where all coordinates of $\mathbf{w}(\cdot)$, $\mathbf{a}(\cdot)$, and $\mathbf{d}(\cdot)$ are differentiable. We have, by the chain rule of differentiation,

$$\begin{aligned} \frac{d}{dt} L(\mathbf{w}(t)) =& \frac{d}{dt}\sqrt{\sum_{k=0}^{c_{\mathcal{K}}} \mathbf{w}_k(t)^2} \\ =& \|\mathbf{w}(t)\|_2^{-1/2} \left(\sum_{k=0}^{c_{\mathcal{K}}} \mathbf{w}_k(t)\dot{\mathbf{w}}_k(t)\right) \\ =& \|\mathbf{w}(t)\|_2^{-1/2} \left(\mathbf{w}_0(t)\dot{\mathbf{w}}_0(t) + \sum_{k \in \mathcal{K}} \mathbf{w}_k(t)\dot{\mathbf{w}}_k(t)\right). \end{aligned} \tag{189}$$

We next inspect separately at the two terms in the parentheses in Eq. (189). For the first term, corresponding to the workload in the Preparation stage, we have

$$\mathbf{w}_0(t)\dot{\mathbf{w}}_0(t) = \mathbf{w}_0(t)(\dot{\mathbf{a}}_0(t) - \dot{\mathbf{d}}_0(t)) \overset{(a)}{=} \mathbf{w}_0(t)(n^P - mq^P) = -\mathbf{w}_0(t)\tilde{c}, \tag{190}$$

where we define $\tilde{c} \triangleq mq^P - n^P > 0$, and step (a) follows from the definition of a fluid solution.

We now analyze the second term, which corresponds to the workloads in the Adaptive stage. Recall that $\mathcal{N}_h$ is the set of vectors satisfying Eqs. (26) through (28). Fix $c_2 > 0$. We will define the following linear program, which we refer to as LP2:

$$\text{minimize} \quad m \tag{191}$$

$$\text{s.t.} \quad (1 + \ln^{-1}(1/\delta)) \sum_{h \in \mathcal{H}} n_{h,k}\pi_h^P \leq r_k m q^A - c_2, \quad \forall k \in \mathcal{K}, \tag{192}$$

$$\{n_{h,k}\}_{k \in \mathcal{K}} \in \mathcal{N}_h, \quad \forall h \in \mathcal{H}, \tag{193}$$

We will denote by m_2^* the optimal value of LP2. For the remainder of the proof, we will assume that

$$mq^A > m_2^*, \tag{194}$$

and demonstrate that (1) Eq. (71) holds whenever the above inequality is true, and (2) the value of m_2^* is not far from the optimal solution to FLP.

From the definition of a fluid solution, we know that

$$\dot{\mathbf{d}}_k(t) = r_k mq^A, \quad \text{if } \mathbf{w}_k(t) > 0, \tag{195}$$

and

$$\{\dot{\mathbf{a}}_k(t)\}_{k\in\mathcal{K}} \leq (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P \mathcal{N}_h^*(\mathbf{w}(t)). \tag{196}$$

For concreteness, we can write

$$\dot{\mathbf{a}}_k(t) = (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P n_{h,k}^*, \quad k \in \mathcal{K}, \tag{197}$$

for some $\{n_{h,k}^*\}_{h\in\mathcal{H},k\in\mathcal{K}}$, where $\{n_{h,k}^*\}_{k\in\mathcal{K}} \leq \mathcal{N}_h^*(\mathbf{w}(t))$ for all $h \in \mathcal{H}$. We have that

$$\begin{aligned}
\sum_{k\in\mathcal{K}} \mathbf{w}_k(t)\dot{\mathbf{w}}_k(t) &= \sum_{k\in\mathcal{K}} \mathbf{w}_k(t)(\dot{\mathbf{a}}_k(t) - \dot{\mathbf{d}}_k(t)) \\
&= \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) \left[(1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P n_{h,k}^* - r_k mq^A\right] \\
&= \sum_{h\in\mathcal{H}} \pi_h^P \left[(1 + \ln^{-1}(1/\delta)) \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n_{h,k}^* - \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) r_k mq^A\right],
\end{aligned} \tag{198}$$

where the last step follows from the fact that $\sum_h \pi_h^P = 1$. Because $mq^A > m_2^*$, the value mq^A must belong to some feasible solution of LP2. By the first constraint of LP2 in Eq. (192), this implies that there exists $\{n_{h,k}\}_{h\in\mathcal{H},k\in\mathcal{K}}$, where $\{n_{h,k}\}_{k\in\mathcal{K}} \in \mathcal{N}_h$ for all $h \in \mathcal{H}$, such that

$$r_k mq^A \geq c_2 + (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P n_{h,k}, \quad \forall k \in \mathcal{K}. \tag{199}$$

We thus have that

$$\begin{aligned}
\sum_{k\in\mathcal{K}} \mathbf{w}_k(t) r_k mq^A \geq& \sum_{k\in\mathcal{K}} \mathbf{w}_k(t)\big(c_2 + (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P n_{h,k}\big) \\
=& c_2 \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) + (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P \left(\sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n_{h,k}\right).
\end{aligned} \tag{200}$$

Substituting the above inequality into Eq. (198), we have that

$$\begin{aligned}
\sum_{k\in\mathcal{K}} \mathbf{w}_k(t)\dot{\mathbf{w}}_k(t) \leq& - c_2 \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) + (1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \pi_h^P \left(\sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n_{h,k}^* - \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n_{h,k}\right) \\
\leq& - c_2 \sum_{k\in\mathcal{K}} \mathbf{w}_k(t),
\end{aligned} \tag{201}$$

where the last inequality follows from the fact that $\{n^*_{h,k}\}_{k\in\mathcal{K}} \leq \mathcal{N}^*_h(\mathbf{w}(t))$, and hence

$$\sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n^*_{h,k} \leq \min_{\{n_k\}_{k\in\mathcal{K}}} \sum_{k\in\mathcal{K}} \mathbf{w}_k(t) n_k. \tag{202}$$

Substituting Eqs. (190) and (201) into Eq. (189), we have that

$$\begin{aligned}
\frac{d}{dt}L(\mathbf{w}(t)) =& \|\mathbf{w}(t)\|_2^{-1/2}\left(\mathbf{w}_0(t)\dot{\mathbf{w}}_0(t) + \sum_{k\in\mathcal{K}} \mathbf{w}_k(t)\dot{\mathbf{w}}_k(t)\right) \\
\leq& -\min\{\tilde{c}, c_2\}\|\mathbf{w}(t)\|_2^{-1/2}\left(\sum_{k=0}^{c_{\mathcal{K}}+1} \mathbf{w}_k(t)\right) \\
\overset{(a)}{\leq}& -\min\{\tilde{c}, c_2\}\|\mathbf{w}(t)\|_2^{-1/2}\|\mathbf{w}(t)\|_2 \\
=& -\min\{\tilde{c}, c_2\}\sqrt{L(\mathbf{w}(t))}
\end{aligned} \tag{203}$$

where in step (a) we used the elementary inequality: $\sum_{i=1}^n x_i \geq \sqrt{\sum_{i=1}^n x_i^2}$, $\forall (x_1,\dots,x_n) \in \mathbb{R}^n_+$.

Recall that all coordinates of $\mathbf{w}(\cdot), \mathbf{a}(\cdot)$ and $\mathbf{d}(\cdot)$ are differentiable for almost all $t \in (0,T)$. Eq. (203) thus implies that $L(\mathbf{w}(t))$ is strictly decreasing whenever $\mathbf{w}(t) \neq 0$. It follows that if $L(\mathbf{w}(0)) = 1$, then, for every $t \in (0,T)$, either $L(\mathbf{w}(t)) \leq 1/4$, or $L(\mathbf{w}(s)) > 1/4$ for all $s \in [0,t]$. In the former case, Eq. (71) is trivially satisfied. In the latter case, we have

$$L(\mathbf{w}(t)) \leq 1 - \min\{\tilde{c}, c_2\}\int_1^t \sqrt{L(\mathbf{w}(s))}ds \leq 1 - \frac{\min\{\tilde{c}, c_2\}}{2}t. \tag{204}$$

Setting the right-hand side of the above equation to $1/4$, we conclude that if

$$\epsilon' = \frac{3}{4} \quad \text{and} \quad \tau = \frac{3}{2\min\{\tilde{c}, c_2\}}, \tag{205}$$

then

$$L(\mathbf{w}(\tau)) \leq 1 - \epsilon', \qquad \text{whenever} \quad L(\mathbf{w}(0)) = 1. \tag{206}$$

Therefore, we have shown that the contraction property of Eq. (71) holds, whenever

$$mq^P > n^P, \text{ and } mq^A > m^*_2. \tag{207}$$

To complete the proof, it remains to relate the optimal value of LP2, m^*_2, to that of the the optimal value of FLP. This is accomplished in the following lemma, whose proof is given in Appendix A.12.

Lemma A.7 *Let m^*_F and m^*_2 be the optimal values of* FLP *and* LP2*, respectively, and c_2 be defined as in* LP2*. For every $\epsilon > 0$, there exist $c' > 0$, such that if $c_2 < c'$, then*

$$m^*_2 q^A \leq \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)}\right)\left(1 + \epsilon^P \frac{c_{\mathcal{H}} c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}}\right)(1 + \ln^{-1}(1/\delta)) m^*_F + \epsilon. \tag{208}$$

The proof of Proposition 7.11 is completed by recalling from Proposition 7.2 that

$$\epsilon^P \leq 2c_{\mathcal{H}}/\ln(1/\delta), \tag{209}$$

and by choosing a sufficiently small ϵ in Lemma A.7. □

A.10 Proof of Proposition 8.1

Proof. Similar to the Markovian state presentation given in Section 7.2 for the Preparation and Adaptive, it is not difficult to show that by incorporating the inspection history of jobs in the Residual stage, the dynamics of the overall system (across all three stages) can be described by an irreducible Markov process. Since the first two stages are assumed to be stable, we now look at the macro-level arrival primitive associated with the Residual stage. The arrivals to the Residual stage are jobs leaving the Adaptive stage who have failed to meet the criterion for exiting the system in Eq. (32). If the Preparation and Adaptive stages are stable, the job arrival process to the Residual stage is generated according to a Markov modulated Poisson process (MMPP), modulated by a stable countable-state Markov process that corresponds to the dynamics of Preparation and Adaptive stages, $\{\mathcal{I}(t), \{Y_i(t)\}_{i\in\mathcal{I}(t)}\}_{t\in\mathbb{R}_+}$, defined in Section 7.2. Using an elementary application of the Foster-Lyapunov criteria, it is thus not difficult to show that the Markov process associated with the overall system is positive recurrent whenever the following rate condition holds: the average rate of inspections from the MMPP should be *less* than the service rate of the Residual stage. In the remainder of the proof, we verify that this is indeed the case.

Recall that $\widehat{H}_i^P$ is the ML estimator job i obtains from the inspections in the Preparation stage. We say that job i is "good", if $\widehat{H}_i^P$ happens to be equal to the true label, H_i, and "bad", otherwise. The next result states that most of the good jobs will be able to depart from the system after being processed in the Adaptive stage, and hence not enter the Residual stage. Denote by $S_i^A(h,l)$ the value of $S_{i,t}(h,l)$ when job i exits the Adaptive stage, and define the event

$$\mathcal{B}^A = \{\exists h' \in \mathcal{H}, \text{ s.t. } S_i^A(h',l) \geq \ln(2c_{\mathcal{H}}/\delta), \forall l \in \mathcal{H}, l \neq h'\}. \tag{210}$$

We have the following lemma, whose proof is given in Appendix A.13.

Lemma A.8 *Fix $i \in \mathbb{N}$. There exists $\delta_0 > 0$, independent of π, such that*

$$\mathbb{P}\left(\mathcal{B}^A \,\middle|\, \widehat{H}_i^P = H_i = h\right) \geq 1 - c_{\mathcal{H}} \ln^{-1}(1/\delta), \quad \forall h \in \mathcal{H}, \tag{211}$$

for all $\delta \in (0, \delta_0)$.

By Proposition 7.2, we have that at least a $(1-2c_{\mathcal{H}}\ln^{-1}(1/\delta))$ fraction of all jobs are good when exiting the Preparation stage. By Lemma A.8, each one of these good jobs has a probability of at most $c_{\mathcal{H}}\ln^{-1}(1/\delta)$ for entering the Residual stage. By assuming that *all* bad jobs eventually enter the Residual stage, the above reasoning shows that the arrival rate of jobs to the Residual stage, λ^R, satisfies

$$\lambda^R \leq 2c_{\mathcal{H}}\ln^{-1}(1/\delta) + c_{\mathcal{H}}\ln^{-1}(1/\delta) = 3c_{\mathcal{H}}\ln^{-1}(1/\delta). \tag{212}$$

In terms of service speed, recall that each job requires n^R inspections in the Residual stage, and the experts visit the Residual stage at the rate of mq^R. Hence, by Eq. (212), it suffices to have

$$\begin{aligned} mq^R >& 3c_{\mathcal{H}}\ln^{-1}(1/\delta)n^R \\ =& 3c_{\mathcal{H}}\ln^{-1}(1/\delta)\zeta_0\ln(4c_{\mathcal{H}}/\delta) \\ =& 3c_{\mathcal{H}}\zeta_0(1+\ln(4c_{\mathcal{H}})\ln^{-1}(1/\delta)). \end{aligned} \tag{213}$$

This proves Proposition 8.1. □

A.11 Proof of Proposition 9.1

Proof. Recall that there are only two ways through which a job could depart from the system: either from the Adaptive stage, or the Residual stage. Fix $i \in \mathbb{N}$ and $h \in \mathcal{H}$. Denote by $S_i^A(h,l)$ the value of $S_{i,t}(h,l)$ at the time job i completes all inspections in the Adaptive stage, and by $\widehat{H}_i^R$ the maximum likelihood estimator of H_i using only the inspections in the Residual stage.

Recall the event $\mathcal{B}^A$ defined in Eq. (210) in Appendix A.10, i.e., whether job i will go through the Residual stage, and denote by $\overline{\mathcal{B}^A}$ its compliment. We have that:

$$\begin{aligned}\mathbb{P}(\widehat{H}_i \neq h \,\big|\, H_i = h) =& \mathbb{P}(\widehat{H}_i \neq h, \mathcal{B}^A \,\big|\, H_i = h) + \mathbb{P}(\widehat{H}_i \neq h, \overline{\mathcal{B}^A} \,\big|\, H_i = h) \\ =& \mathbb{P}(\widehat{H}_i \neq h, \mathcal{B}^A \,\big|\, H_i = h) + \mathbb{P}(\widehat{H}_i \neq h \,\big|\, \overline{\mathcal{B}^A}, H_i = h)\mathbb{P}(\overline{\mathcal{B}^A} \,\big|\, H_i = h) \\ \overset{(a)}{=}& \mathbb{P}(\widehat{H}_i \neq h, \mathcal{B}^A \,\big|\, H_i = h) + \mathbb{P}(\widehat{H}_i^R \neq h \,\big|\, \overline{\mathcal{B}^A}, H_i = h)\mathbb{P}(\overline{\mathcal{B}^A} \,\big|\, H_i = h) \\ \leq& \mathbb{P}(\widehat{H}_i \neq h, \mathcal{B}^A \,\big|\, H_i = h) + \mathbb{P}(\widehat{H}_i^R \neq h \,\big|\, \overline{\mathcal{B}^A}, H_i = h), \end{aligned} \tag{214}$$

where step (a) follows from the definition of the inspection policy, where a job will be sent to the Residual stage if and only if the event $\mathcal{B}^A$ does not occur. The two terms on the right-hand-side of Eq. (214) correspond to the classifications made by the Adaptive and Residual stages, respectively. By noting that the total number of inspections received by job i by the time it exits the Adaptive stage is a stopping time with respect to the past inspections, applying Lemma A.1, with $x = \ln(2c_{\mathcal{H}}/\delta)$, we have that

$$\mathbb{P}(\widehat{H}_i \neq h, \mathcal{B}^A \,\big|\, H_i = h) \leq \delta/2. \tag{215}$$

Now suppose that job i exits the system after having gone through the Residual stage. Recall that by construction, the Residual stage employs an inspection procedure that is the same as the Preparation stage, except for having a larger number of inspections per job. Therefore, using an essentially identical line of arguments to that of Proposition 7.2, and, in particular, to the portion of the proof leading to Eq. (142), by replacing the quantity $\ln\ln(1/\delta)$ with $\ln(2c_{\mathcal{H}}/\delta)$, we can show that

$$\mathbb{P}(\widehat{H}_i^R \neq h \,\big|\, \overline{\mathcal{B}^A}, H_i = h) \leq \delta/2. \tag{216}$$

Substituting Eqs. (215) and (216) into Eq. (214) completes the proof of Proposition 9.1. □

A.12 Proof of Lemma A.7

Proof.

Let $(m_F^*, \{n_{h,k}^*\})$ be an optimal solution to FLP that satisfies the conditions stated in Lemma 6.3. Define

$$\phi_\delta = 1 + (\ln(2c_{\mathcal{H}}) + g_\delta)\ln^{-1}(1/\delta), \tag{217}$$

and let

$$\tilde{n}_{h,k} = n_{h,k}^* \phi_\delta. \tag{218}$$

Recall from the definition in Eq. (30) that

$$v_\delta = 2c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)[1 + (\ln(2c_{\mathcal{H}}) + g_\delta)\ln^{-1}(1/\delta)] = 2c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\phi_\delta, \tag{219}$$

which implies, by Eq. (39) of Lemma 6.3, that $\sum_{k\in\mathcal{K}} \tilde{n}_{h,k} \leq v_\delta$ for all $h \in \mathcal{H}$. We thus conclude that the variables $\{\tilde{n}_{h,k}\}$ satisfy the second set of constraints of LP2, i.e.,

$$\{\tilde{n}_{h,k}\} \in \mathcal{N}_h, \quad \forall h \in \mathcal{H}. \tag{220}$$

We now derive a sufficient condition on $\tilde{m}$ so that $(\tilde{m}, \{\tilde{n}_{h,k}\})$ is a feasible solution to LP2. Fixing $k \in \mathcal{K}$, we have that

$$\begin{aligned}\sum_{h\in\mathcal{H}} \tilde{n}_{h,k}\pi_h^P =& \phi_\delta \sum_{h\in\mathcal{H}} n^*_{h,k}\pi_h^P \\ &\overset{(a)}{\leq} \phi_\delta \left(\sum_{h\in\mathcal{H}} n^*_{h,k}\pi_h + \epsilon^P \sum_{h\in\mathcal{H}} n^*_{h,k} \right) \\ &\overset{(b)}{\leq} \phi_\delta \left(\sum_{h\in\mathcal{H}} n^*_{h,k}\pi_h + \epsilon^P c_{\mathcal{H}} c_{\mathcal{K}} \underline{d}^{-1} \ln(1/\delta) \right) \\ &\overset{(c)}{\leq} \phi_\delta \left(\sum_{h\in\mathcal{H}} n^*_{h,k}\pi_h + \epsilon^P c_{\mathcal{H}} c_{\mathcal{K}} \underline{d}^{-1} \overline{d} m^*_F \right) \\ &\overset{(d)}{\leq} \phi_\delta \left(m^*_F r_k + \epsilon^P c_{\mathcal{H}} c_{\mathcal{K}} \underline{d}^{-1} \overline{d} m^*_F \right).\end{aligned} \tag{221}$$

Step (a) follows from Proposition 7.2, (b) and (c) from Lemma 6.3, and (d) from the constraint of FLP in Eq. (35). To satisfy the first constraint of LP2, it suffices to let the right-hand side of Eq. (221) satisfy

$$\phi_\delta \left(m^*_F r_k + \epsilon^P c_{\mathcal{H}} c_{\mathcal{K}} \underline{d}^{-1} \overline{d} m^*_F \right) \leq \frac{r_k q^A \tilde{m} - c_2}{1 + \ln^{-1}(1/\delta)}. \tag{222}$$

That is, if

$$\tilde{m} q^A \geq m^*_F \phi_\delta \left(1 + \epsilon^P \frac{c_{\mathcal{H}} c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}} \right) (1 + \ln^{-1}(1/\delta)) + \frac{c_2}{\underline{r}}, \tag{223}$$

then

$$(1 + \ln^{-1}(1/\delta)) \sum_{h\in\mathcal{H}} \tilde{n}_{h,k}\pi_h^P \leq r_k q^A \tilde{m} - c_2, \quad \forall k \in \mathcal{K}, \tag{224}$$

which, in light of the fact that $\{\tilde{n}_{h,k}\}_{k\in\mathcal{K}} \in \mathcal{N}_h$ for all $h \in \mathcal{H}$, further implies that $(\tilde{m}, \{\tilde{n}_{h,k}\})$ is a feasible solution of LP2. Therefore, we conclude that, for all $c_2 > 0$,

$$\begin{aligned} m^*_2 q^A \leq& \phi_\delta \left(1 + \epsilon^P \frac{c_{\mathcal{H}} c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}} \right) (1 + \ln^{-1}(1/\delta)) m^*_F + \frac{c_2}{\underline{r}} \\ =& \left(1 + \frac{\ln(2c_{\mathcal{H}}) + g_\delta}{\ln(1/\delta)} \right) \left(1 + \epsilon^P \frac{c_{\mathcal{H}} c_{\mathcal{K}} \overline{d}}{\underline{d}\,\underline{r}} \right) (1 + \ln^{-1}(1/\delta)) m^*_F + \frac{c_2}{\underline{r}}, \end{aligned} \tag{225}$$

We complete the proof of Lemma A.7 by letting $c' = \epsilon \underline{r}$. □

A.13 Proof of Lemma A.8

Proof. The proof follows similar steps as those in the proof of Lemma A.2. Fix $h \in \mathcal{H}$ and $\{\lambda_{i,k}\}_{k\in\mathcal{K}} \in \mathbb{Z}_+^{c_{\mathcal{K}}}$. Denote by $\mathcal{B}$ the event

$$\mathcal{B} = \{\widehat{H}_i^P = H_i = h, \{\Lambda_{i,k}\}_{k\in\mathcal{K}} = \{\lambda_{i,k}\}_{k\in\mathcal{K}}\} \tag{226}$$

For the remainder of the proof, we will assume that $h \in \mathcal{H}$ and $\{\lambda_{i,k}\}_{k\in\mathcal{K}}$ are such that $\mathbb{P}(\mathcal{B}) > 0$. We will index all inspections received by job i in the Adaptive stage in an arbitrary fashion, and denote by K_n the type of the expert who performed the nth inspection for job i, and by $Z_n(\cdot,\cdot,K_n)$ the corresponding log-likelihood ratio. Define $\Lambda_i^0 = \sum_{k\in\mathcal{K}} \Lambda_{i,k}$. For all $l \in \mathcal{H}, l \neq h$, let

$$M_n^l = \sum_{s=1}^{n\vee\Lambda_i^0} Z_s(h,l,K_s) - D(h,l,K_s), \quad n \in \mathbb{N}, \tag{227}$$

and $M_0^l = 0$. It is not difficult to see that, conditional on the true label of job i being h, the summands in the above equation have zero mean and are independent. Therefore, one can verify that $\{M_n^l \,|\, \mathcal{B}\}_{n\in\mathbb{Z}_+}$ is a martingale.

Define

$$\phi_\delta = 1 + (\ln(2c_{\mathcal{H}}) + g_\delta)\ln^{-1}(1/\delta), \tag{228}$$

By the construction of $\{\Lambda_{i,k}\}_{k\in\mathcal{K}}$ (Eq. (31)), we know that

$$\Lambda_i^0 \leq \sum_{k\in\mathcal{K}} n_k \leq v_\delta = 2\phi_\delta c_{\mathcal{K}} \underline{d}^{-1}\ln(1/\delta). \tag{229}$$

By Eq. (26), we know that if

$$M_{\Lambda_i^0}^l - M_0^l > -g_\delta + c_{\mathcal{K}}\overline{d}, \tag{230}$$

then

$$S_i^A(h,l) > \ln(2c_{\mathcal{H}}/\delta), \tag{231}$$

where the term $c_{\mathcal{K}}\overline{d}$ captures the potential discrepency induced by the rounding in Eq. (31). We have that

$$\begin{aligned}\mathbb{P}\left(S_i^A(h,l) \leq \ln(2c_{\mathcal{H}}/\delta) \,\middle|\, \mathcal{B}\right) =& \mathbb{P}\left(M_{\Lambda_i^0}^l - M_0^l \leq -(g_\delta^A - c_{\mathcal{K}}\overline{d}) \,\middle|\, \mathcal{B}\right)\\ &\overset{(a)}{\leq} \exp\left(-\frac{(g_\delta - c_{\mathcal{K}}\overline{d})^2}{4\overline{z}^2\Lambda_i^0}\right)\\ &\overset{(b)}{\leq} \exp\left(-\frac{(g_\delta)^2 - 2g_\delta c_{\mathcal{K}}\overline{d}}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta\ln(1/\delta)}\right).\end{aligned} \tag{232}$$

Step (a) follows from the Azuma-Hoeffding's Inequality (Lemma C.1 in Appendix C), and noting that $|M_n^l - M_{n-1}^l| \leq 2\overline{z}$ for all $n \in \mathbb{N}$. Step (b) follows from Eq. (229).

Recall that $g_\delta = 3\overline{z}\sqrt{c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\ln\ln(1/\delta)}$. The exponent in Eq. (232) can be further expanded as:

$$\begin{aligned}-\frac{(g_\delta)^2 - 2g_\delta c_{\mathcal{K}}\overline{d}}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta\ln(1/\delta)} &= -\frac{9\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\ln(1/\delta)\ln\ln(1/\delta)}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta\ln(1/\delta)} + \frac{6\overline{z}\underline{d}^{-1/2}c_{\mathcal{K}}^{3/2}\overline{d}\sqrt{\ln(1/\delta)\ln\ln(1/\delta)}}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta\ln(1/\delta)}\\ &= -\frac{9\ln\ln(1/\delta)}{8\phi_\delta} + \frac{3\underline{d}^{-3/2}c_{\mathcal{K}}^{1/2}\overline{d}}{4\overline{z}}\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}.\end{aligned} \tag{233}$$

As $\delta \downarrow 0$, we have that $\phi_\delta = 1 + (\ln(2c_{\mathcal{H}}) + g_\delta)\ln^{-1}(1/\delta) \to 1$, and $\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}} \to 0$. Therefore, the above expression yields[9]

$$-\frac{(g_\delta)^2 - 2g_\delta c_{\mathcal{K}}\overline{d}}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta \ln(1/\delta)} \sim -\frac{9\ln\ln(1/\delta)}{8}, \quad \text{as } \delta \to 0. \tag{234}$$

Substituting Eq. (234) into Eq. (232), we know that there exists $\delta_0 > 0$, independent of the prior distribution, π, and the choice of l and h, such that for all $\delta \in (0, \delta_0)$,

$$\begin{aligned}
\mathbb{P}\left(S_i^A(h,l) \leq \ln(2c_{\mathcal{H}}/\delta) \,\middle|\, \mathcal{B}\right) &\leq \exp\left(-\frac{(g_\delta)^2 - 2g_\delta c_{\mathcal{K}}\overline{d}}{8\overline{z}^2 c_{\mathcal{K}}\underline{d}^{-1}\phi_\delta \ln(1/\delta)}\right) \\
&= \exp\left(-\frac{9\ln\ln(1/\delta)}{8\phi_\delta} + \frac{3\underline{d}^{-3/2}c_{\mathcal{K}}^{1/2}\overline{d}}{4\overline{z}}\sqrt{\frac{\ln\ln(1/\delta)}{\ln(1/\delta)}}\right) \\
&\leq \exp(-\ln\ln(1/\delta)) \\
&= \ln^{-1}(1/\delta).
\end{aligned} \tag{235}$$

Since the above inequalities hold for all $h, l \in \mathcal{H}$, $l \neq h$, and $\{\lambda_{i,k}\}_{k\in\mathcal{K}}$, we can apply a union bound over l and conclude that, for all $h \in \mathcal{H}$,

$$\mathbb{P}(\mathcal{B}^A \,|\, \widehat{H}_i^P = H_i = h) \geq \mathbb{P}\left(S_i^A(h,l) \geq \ln(2c_{\mathcal{H}}/\delta),\, \forall l \neq h \,\middle|\, \widehat{H}_i^P = H_i = h\right) \geq 1 - c_{\mathcal{H}}\ln^{-1}(1/\delta). \tag{236}$$

This completes the proof of Lemma A.8. □

B Some Properties of $L(\cdot)$

The following lemma states some basic properties of the Lyapunov function, $L(\cdot)$.

Lemma B.1 *The function $L(\mathbf{w}) = \|\mathbf{w}\|_2$ admits the following properties: (1) L is continuous on $\mathbb{R}_+^{c_{\mathcal{K}}+1}$; (2) For all $c \in \mathbb{R}_+$ and $\mathbf{w} \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$, we have $L(c\mathbf{w}) = cL(\mathbf{w})$; (3) There exists positive constants, $\alpha_1, \alpha_2 > 0$, such that for all $\mathbf{w} \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$,*

$$\alpha_1\|\mathbf{w}\|_\infty \leq L(\mathbf{w}) \leq \alpha_2\|\mathbf{w}\|_\infty, \tag{237}$$

where $\|\cdot\|_\infty$ is the l_∞ norm, $\|\mathbf{w}\|_\infty = \max_{k=0,\ldots,c_{\mathcal{K}}+1}|\mathbf{w}_k|$.

Proof. The first two properties follow directly from the definition of $L(\cdot)$. For Eq. (237), we note that for all $\mathbf{w} \in \mathbb{R}_+^{c_{\mathcal{K}}+1}$,

$$\begin{aligned}
L(\mathbf{w}) &\geq \sqrt{\|\mathbf{w}\|_\infty^2} = \|\mathbf{w}\|_\infty, \\
L(\mathbf{w}) &\leq \sqrt{(c_{\mathcal{K}}+1)\|\mathbf{w}\|_\infty^2} = \sqrt{c_{\mathcal{K}}+1}\|\mathbf{w}\|_\infty.
\end{aligned}$$

Hence, Eq. (237) holds for $\alpha_1 = 1$ and $\alpha_2 = \sqrt{c_{\mathcal{K}}+1}$. □

[9] The notation $f(x) \sim g(x)$ means that $\lim_{x\to 0}\frac{f(x)}{g(x)} = 1$.

C The Azuma-Hoeffding Inequality

We will make use of the standard Azuma-Hoeffding inequality (Section 12.2 of Grimmett and Stirzaker (2001)), which we state below for completeness.

Lemma C.1 (Azuma-Hoeffding Inequality) *Let $\{M_n\}_{n\in\mathbb{Z}_+}$ be a martingale, and suppose that there exists $a > 0$, such that $|M_n - M_{n-1}| \le a$ for all $n \in \mathbb{Z}_+$. Then*

$$\mathbb{P}(M_n - M_0 \le -x) \le \exp\left(-\frac{x^2}{2na^2}\right), \quad \forall x \in \mathbb{R}_+. \tag{238}$$

D A Heuristic Policy

We discuss in this section a simple heuristic inspection policy, which essentially condenses the three-stage architecture of Figure 1 into only one that resembles the Adaptive stage. We expect the policy to be substantially easier to implement than the one presented in Section 5. Nevertheless, we have not been able to establish its resource efficiency rigorously, though we discuss some indication that it may be resource efficient as well.

For simplicity, we assume that $D(h,l,k) > 0$ for all $k \in \mathcal{K}, h \neq l$, i.e., every expert is able to distinguish any two job labels to some degree. Denote by $\widehat{H}_{i,t}$ the ML estimator of job i's type at time t, and we assume that the value of $\widehat{H}_{i,t}$ at the time when job i has just arrived to the system is set to a value drawn uniformly at random from $\mathcal{H}$. Denote by $\mathcal{I}_t$ the indices of all jobs that are currently in the system at time t, and denote by $\mathcal{I}_{h,t}$ the subset containing those jobs whose ML estimator is equal to h:

$$\mathcal{I}_{h,t} = \{i \in \mathcal{I}_t : \widehat{H}_{i,t} = h\}. \tag{239}$$

Define $W_{i,t}(h,l)$ as the *residual workload* associated with job i:

$$W_{i,t}(h,l) = \left(\ln(c_{\mathcal{H}}/\delta) - S_{i,t}(h,l)\right)^+, \quad h,l \in \mathcal{H}, \tag{240}$$

and define the *aggregate residual workload* associated with job type h:

$$\overline{W}_t(h,l) = \sum_{i\in\mathcal{I}_{h,t}} W_{i,t}(h,l), \quad h,l \in \mathcal{H}. \tag{241}$$

The heuristic policy operates as follows.

Departure rule. Job i departs from the system as soon as there exists $h \in \mathcal{H}$, such that

$$\max_{l\in\mathcal{H}, l\neq h} W_{i,t}(h,l) = 0. \tag{242}$$

Expert actions. Suppose that an expert of type k becomes available at time t. If there is no job in the system, she goes on a vacation. Otherwise, she inspects the oldest job from the set $\mathcal{I}_{h^*,t}$, where

$$h^* \in \arg\max_{h\in\mathcal{H}} \sum_{l\in\mathcal{H}, l\neq h} D(h,l,k)\overline{W}_t(h,l), \tag{243}$$

with ties broken arbitrarily.

This heuristic policy has a number of nice features. As is evident from the above description, the policy is much simpler than the one given in Section 5, and it does not require solving a linear optimization problem as a sub-routine. Also, the departure criterion ensures that the event $\mathcal{G}_x$ in Lemma 4.1 always occurs for every job that leaves the system, so the policy is δ-accurate whenever it is stable. There is also some indication that this heuristic policy might be resource efficient. For a moment, let us suppose that all ML estimators, $\widehat{H}_{i,t}$, are correct. In this case, $\mathcal{I}_{h,t}$ is the set of jobs whose true label is h, and Eq. (243) can be thought of as a max-weight-like procedure, where $D(h,l,k)$ corresponds to the expected decrease a single inspection can incur on the aggregate residual workload $\overline{W}_t(h,l)$. Since max-weight scheduling policies (e.g., Tassiulas and Ephremides (1992)) are known to achieve the maximum stability region in many queueing systems, one may expect that this max-weight-like property would make our heuristic policy resource efficient as well.

Unfortunately, there appear to be two intrinsic characteristics of the policy that make it difficult to rigorously establish resource efficiency. Firstly, the ML estimators $\widehat{H}_{i,t}$ can be *incorrect*, especially when a job has received only a small number of inspections, and therefore the vector $\{\overline{W}_t(h,l)\}_{h,l\in\mathcal{H}}$ does not exactly capture the "true" workload of the system. Secondly, even assuming that the ML estimators are correct, a synchronization issue can prevent a type-k expert's inspection from actually contributing an expected $D(h,l,k)$ units of decrease in $\overline{W}_t(h,l)$ for all $l \neq h$. This is because, for some job i, $W_{i,t}(h,l)$ along some coordinate l could reach zero much earlier than other coordinates $l' \neq l$, during which period the expected decrease in $\overline{W}_t(h,l)$ from inspecting job i would be zero, instead of $D(h,l,k)$, due to the capping at zero of Eq. (242). Both of the above issues make it difficult to bound the drift in a Lyapunov function in the course of proving stability. However, there is hope that they could be mitigated as the accuracy parameter, δ, tends to 0. In this regime, the number of inspections needed for each job tends to infinity, and hence most jobs in the system have already received a large number of inspections, implying that they likely have the correct ML estimators as well. Similarly, since the distortion of drift only occurs at the "boundary" where the residual workload of a job is small, it will likely have much less impact as the number of inspections per job grows.

E Glossary of Frequently Used Symbols

$c_{\mathcal{H}}$	cardinality of $\mathcal{H}$
$c_{\mathcal{K}}$	cardinality of $\mathcal{K}$
$\mathcal{E}$	set of experts
δ	classification error, accuracy requirement
$D(h, l, k)$	KL-divergence from $p(h, k, \cdot)$ to $p(l, k, \cdot)$
$\overline{d}$	maximum value among $D(h, l, k)$
$\underline{d}$	minimum value among non-zero $D(h, l, k)$
$\mathcal{H}$	set of job labels
H_i	true label of job i
$\hat{H}_i$	classification for job i's label
$\mathcal{K}$	set of expert types
λ_0	job arrival rate
μ_k	inspection rate of type-k experts
m	system size; number of experts
m^*	optimal/minimum system size
$p(h, k, \cdot)$	outcome distribution when a job with true label h is inspected by a type-k expert
π	prior distribution of job labels
ρ	expert mixture
$S_{i,t}$	cumulative log-likelihood ratios associated with job i by time t
$\mathbf{W}$	workload process
$X_{i,j}$	outcome of the jth inspection performed on job i
$\mathcal{X}$	set of all possible outcomes
$Z_{i,j}$	log-likelihood ratios associated with the jth inspection on job i
$\overline{z}$	maximum log-likelihood ratio