Amortized Analysis on Asynchronous Gradient Descent

Yun Kuen Cheung[*]
University of Vienna

Richard Cole
Courant Institute, NYU

Abstract

Gradient descent is an important class of iterative algorithms for minimizing convex functions. Classically, gradient descent has been a sequential and synchronous process. Distributed and asynchronous variants of gradient descent have been studied since the 1980s, and they have been experiencing a resurgence due to demand from large-scale machine learning problems running on multi-core processors.

We provide a version of asynchronous gradient descent (AGD) in which communication between cores is minimal and for which there is little synchronization overhead. We also propose a new timing model for its analysis. With this model, we give the first amortized analysis of AGD on convex functions. The amortization allows for bad updates (updates that increase the value of the convex function); in contrast, most prior work makes the strong assumption that every update must be significantly improving.

Typically, the step sizes used in AGD are smaller than those used in its synchronous counterpart. We provide a method to determine the step sizes in AGD based on the Hessian entries for the convex function. In certain circumstances, the resulting step sizes are a constant fraction of those used in the corresponding synchronous algorithm, enabling the overall performance of AGD to improve linearly with the number of cores.

We give two applications of our amortized analysis:

- We show that our AGD algorithm can be applied to two classes of problems which have huge problem sizes in applications and consequently can benefit substantially from parallelism. The first class of problems is to solve linear systems $Ap = b$, where the A are symmetric and positive definite matrices. The second class of problems is to minimize convex functions of the form $\sum_{i=1}^{n} f_i(p_i) + \frac{1}{2}\|Ap - b\|^2$, where the f_i are convex differentiable univariate functions.
- We show that a version of asynchronous tatonnement, a simple distributed price update dynamic, converges toward the market equilibrium in Fisher markets with buyers having complementary-CES or Leontief utility functions.

[*]Most of the work done while at Courant Institute, NYU.

1 Introduction

Gradient descent, an important class of iterative algorithms for minimizing convex functions, is a key subroutine in many computational problems. Broadly speaking, gradient descent proceeds by iteratively moving in the direction of the negative gradient of the convex function. Classically, gradient descent is a sequential and synchronous process. Distributed and asynchronous variants have also been studied, starting with the work of Tsitsiklis et al. [17] in the 1980s; more recent results include [2, 3]. Distributed and asynchronous gradient descent has been experiencing a resurgence of attention, particularly in computational learning theory [12, 15], due to recent advances in multi-core parallel processing technology and a strong demand for speeding-up large-scale gradient descent problems via parallelism.

Gradient descent proceeds by repeatedly updating the coordinates of the argument to the convex function. A few key common issues arise in any distributed and asynchronous iterative implementation and their improper handling may lead to performance-destroying overhead costs.

- In some implementations (e.g. [15]), different cores[1] may update the same component. Without proper coordination, the progress made by one core can be overwritten, and if such overwriting persists, in the worst case the system can fail to reach the desired result.

  This difficulty can be avoided by block component descent – each coordinate is updated by exactly one core. This is the approach we use in our Asynchronous Gradient Descent (AGD) algorithm. The approach has been used previously in a round-robin manner [12], but our AGD algorithm does not require the updates to proceed in any particular order.

- The cores need to follow a communication protocol in order to communicate/broadcast their updates. Communication is often relatively slow compared to computation, so reducing the need for communication can lead to a significant improvement in system performance. Also, when there is delay in communication, cores may use outdated information for the next update, which is a critical issue for asynchronous systems.

  One common approach is to assume that the system has *bounded asynchrony*, i.e. the delay in communication is bounded by a positive constant. Typically, there is a need to wait for updates from the other cores, and the bounded asynchrony simply bounds the waiting time. We will use the bounded asynchrony assumption, but our AGD algorithm will have *no waiting*: updates will always be based on the information at hand; bounded asynchrony just guarantees that it is not too dated.

- Often, the computation of one core needs the results computed by another core, implying the computations of the different cores must be in a correct order to ensure correctness and to reduce core waiting time. Typically this is achieved via a synchronization protocol, which often requires that all cores follow a global clock. However, such protocols can be costly and even impractical in some circumstances.

  As we shall see, our AGD algorithm needs essentially no synchronization apart from an initial synchronization to align the starting times of all cores.

[1] These observations apply to any multi-processor system.

Broadly speaking, most prior work follows the asynchrony model proposed in [17], in which time is discretized. Our AGD algorithm allows each core to proceed at its own pace. This allows for varying loads, for different updates having varied costs, for interruptions, and more generally for variations in the completion times of updates. To support this, in our model, time is continuous. To ensure progress, we require that each component be updated at least once in each time unit, but do not impose an upper bound on the frequency of updates. A more formal description of our model will be given in Section 2.

We consider a robust family of AGD algorithms, and using our timing model, we give a new amortized analysis which shows each algorithm converges to the minimal value of the underlying function. Most prior work made the strong assumption that each update yields a significant improvement. Our analysis, however, allows for bad individual updates (updates that increase the value of the convex function), which seem to be unavoidable in general. In our AGD algorithm, every update leads to errors in subsequent gradient measurements at other cores. A natural question to ask is whether such errors can propagate and be persistent and whether they might, in the worst case, prohibit convergence toward a minimal point. Our amortized analysis shows that this will not happen when the step sizes used in the AGD algorithm are suitably bounded. The following observation forms a key part of the analysis: if there is a bad update to one component, it can only be due to some recent good updates to other components, or to chaining of this effect. We use a carefully designed potential function, which *saves* a portion of the gains due to good updates, to pay for the bad updates. The amortized analysis will be presented in Section 3.

Typically the step sizes used in AGD are smaller than those used in its synchronous counterpart. Our AGD algorithm determines the step sizes based on the Hessian of the underlying function. In certain circumstances, the step sizes in our AGD can be a constant fraction of those used in its synchronous counterpart, ensuring that the number of rounds of updates performed by the AGD algorithm is within a constant of the analogous upper bound for the synchronous version. Note that AGD avoids the synchronization costs of its synchronous counterpart, which are a practical concern [15].

Application: Solving Matrix Systems in Parallel We begin by considering two problems in which bad updates are possible in an asynchronous setting. A linear system is the problem of finding $p \in \mathbb{R}^n$ that satisfies $Ap = b$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are the inputs. As is well-known, if A is a symmetric and positive definite matrix, solving the linear system is equivalent to finding the minimum point of a strongly convex function, so our AGD algorithm can be applied.

Nesterov [14] discusses the following class of optimization problems: minimizing convex functions of the form $\sum_{i=1}^n f_i(p_i) + \frac{1}{2}\|Ap - b\|^2$, where the f_i are convex differentiable univariate functions. The size of such problems can be huge in practice, and input/data can be distributed in space and time, so time synchronization is costly and even impractical. One important feature of our AGD algorithm is to allow the use of data that are *variously dated*. As we will see, this hugely reduces the need for synchronization. More details are given in Section 4.

Application: Asynchronous Tatonnement in Fisher Markets We show that an asynchronous tatonnement converges toward the market equilibrium in two classes of Fisher markets.

The concept of a market equilibrium was first proposed by Walras [19]. Walras also proposed an algorithmic approach for finding equilibrium prices, namely to adjust prices by tatonnement: upward if there is too much demand and downward if too little. Since then, the study of market equilibria and tatonnement have received much attention in economics, operations research, and most recently in computer science [1, 18, 8, 16]. Underlying many of these works is the issue of what are plausible price adjustment mechanisms and in what types of markets they attain a market equilibrium.

The tatonnements studied in prior work have mostly been continuous, or discrete and synchronous. Observing that real-world market dynamics are highly distributed and hence presumably asynchronous, Cole and Fleischer [10] initiated the study of asynchronous tatonnement with their *Ongoing market model*, a market model incorporating update dynamics.

Cheung, Cole and Devanur [6] showed that tatonnement is equivalent to gradient descent on a convex function for several classes of Fisher markets, and consequently that a suitable synchronous tatonnement converges toward the market equilibrium in two classes of markets: complementary-CES Fisher markets and Leontief Fisher markets. This equivalence also enables us to apply our amortized analysis to show that the corresponding asynchronous version of tatonnement converges toward the market equilibrium in these two classes of markets. More details are given in Section 5. We note that the tatonnement for Leontief Fisher markets that was analysed in [6] has an unrealistic constraint on the step sizes; our analysis removes that constraint, and works for both synchronous and asynchronous tatonnement.

2 Asynchronous Gradient Descent Model

We consider the following unconstrained optimization problem: given a convex function ϕ: $\mathbb{R}^n \to \mathbb{R}$, find its minimal point. In our model, time, denoted by t, is continuous. The gradient descent process starts at $t = 0$ from an initial point $p^0 = (p_1^0, p_2^0 \cdots, p_n^0)$. For simplicity, we assume that there are n cores, and p_j is updated by the j-th core.[2] After each update, the updating core broadcasts it; the other cores receive the message, possibly with a delay.

Notational Convention When there is an update at time t which updates the value of one or more variables, for each such variable $\Box$, we let both $\Box^{t-}$ and $\Box^t$ denote its value just before the update, and $\Box^{t+}$ its value right after the update.

We define $p^t \equiv p^{t-}$, the *current point* at time t, to comprise the most recently updated values for each coordinate. However, any particular core may have out-of-date values for one or more coordinates, but not too much out-of-date, as we specify next.

Let t_1 and t_2 be the times of successive updates to p_j. Then, at time t_2, the j-th core will have values for each of the other coordinates that were current at time t_1 or later. In other words, the time taken to communicate an update is no larger than $t_2 - t_1$. Effectively, this is the constraint on how much parallelism is possible. Informally speaking, the information which the core holds is at most one "round" out of date w.r.t. its updates. In fact, it seems likely that we could extend our analysis to allow for any fixed constant number of rounds of datedness, but as this would entail a proportionate reduction in the step sizes, it does not seem useful.

[2]If there are fewer cores it suffices to cluster coordinates.

However, there is no requirement that updates occur at a similar rate, although we imagine that this would be the typical case. It may be natural in some settings for coordinates to adjust with different frequencies, e.g. prices of different goods in a broad enough market. Accordingly, we define a rather general update rule, as follows. Each core has the freedom to determine the time at which it updates its coordinate. To proceed, it will be helpful to define the following rectangular subsets of coordinate values.

Definition 1. $\tilde{P}_j^{[t_1,t_2]}(s_j)$ *comprises the rectangular box with* $p_j = s_j$ *and, for* $k \neq j$*, spanning the range of values* p_k *that occur over the time interval* $[t_1, t_2]$.

Let τ_j be the time at which the last update to p_j occurred, and let t be the time of the current update to p_j. To update p_j, the j-th core computes $\nabla_j \phi(\tilde{p})$, where $\tilde{p}$ is an arbitrary point in $\tilde{P}_j^{[\tau_j,t]}\left(p_j^t\right)$. This flexibility allows different coordinates at the j-th core to be *variously dated*, under the constraint that they are all no older than time τ_j. The general form of an update is

$$p_j \leftarrow p_j + F_j(\tilde{p}, \nabla_j \phi(\tilde{p}), t) \cdot (t - \tau_j),$$

where F_j is a function such that $F_j(\tilde{p}, \nabla_j \phi(\tilde{p}), t)$ has the same sign as $-\nabla_j \phi(\tilde{p})$.

The term $t - \tau_j$ is somewhat unusual. It is needed because we impose no bound on the frequency of updates. Without this multiplier, a core, the k-th core say, could perform many updates in the time interval $[\tau_j, t]$, potentially making a cumulatively large update to p_k, which could lead to an unbounded difference between $\nabla_j \phi(\tilde{p})$ and $\nabla_j \phi(p^t)$. This appears to preclude the usual approaches to a proof of convergence, and even calls convergence into question in general. If, in fact, $t - \tau_j = \Theta(1)$ always, then this term can be omitted.

Note that the sign of $F_j(\tilde{p}, \nabla \phi(\tilde{p}), t)$ can be opposite to that of $F_j(p^t, \nabla_j \phi(p^t), t)$; when this occurs, an update will increase the value of ϕ, i.e. we have a bad update!

We do not require any further coordination between the cores. We just require a minimal amount of communication to ensure that the cores know an approximation of the current point so that they can compute a useful gradient.

3 Amortized Analysis

Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a twice-differentiable convex function. Our AGD algorithm solves the problem of finding (or approximating) a minimal point of ϕ, which we denote by p^*. WLOG, we assume that $\phi^* := \phi(p^*) = 0$. We assume that no two updates occur at the same time.[3]

By default, each core possesses the most up-to-date entry for the coordinate it updates. However, due to communication delay, it may have outdated entries for coordinates updated by other cores. Recall that p^t denotes the most up-to-date entries at time t; let $\tilde{p}_k^{j,t}$ denote the entry for p_k that the j-th core possesses at time t. Note that $\tilde{p}^{j,t} \in \tilde{P}_j^{[\tau_j,t]}\left(p_j^t\right)$.

We now consider an update to p_j at time t given by

$$p'_j \leftarrow p_j - \frac{\tilde{g}_j(t)}{\gamma_j^t} \Delta t_j, \tag{1}$$

[3] If two or more updates do occur at the same time, our analysis remains valid by making infinitesimal perturbations to their update times.

where $\tilde{g}_j(t) = \nabla_j \phi(\tilde{p}^{j,t})$, $\Delta t_j = t - \tau_j$, and $1/\gamma_j^t$ is the *step size*, which will be determined by a rule we specify later. We assume that $\Delta t_j \le 1$ always, i.e. two consecutive updates to the same coordinate occur at most one time unit apart. We note that Rule (1) is quite general for it allows both additive and multiplicative updates, depending on the choice of the γ_j^t. As we shall see, our analysis handles applications of both types.

For any $S \subset \mathbb{R}^n$, let $H_{k\ell}(S) := \max_{p' \in S} \left| \frac{\partial^2 \phi}{\partial p_k \partial p_\ell}(p') \right|$. We will use the shorthand $H_{k\ell}^{[t_1,t_2]}(s_\ell)$ for $H_{k\ell}\left(\tilde{P}_\ell^{[t_1,t_2]}(s_\ell)\right)$. In order to show our convergence results, the γ_j^t need to be suitably constrained and the Hessian entries need to be sufficiently bounded. We capture this in our definition of *controlled* γ_j^t and H_{jk}, given right after Theorem 1 below.

Theorem 1. *Suppose that all updates are made according to update rule* (1)*. Let* $\overline{\gamma} = \max_{j,t} \gamma_j^t$*. If the variables* γ_j^t *and* H_{jk} *are controlled, then*

(a) Suppose the set $\{p' \mid \phi(p') \le 2\phi(p^0)\}$ *is bounded with diameter* B*. Let* $M(B) := \Theta(B^2 \overline{\gamma})$*. Then, if* $\phi(p^0) \le M(B)$*,* $\phi(p^t) = O\left(\frac{M(B)}{t}\right)$*; and otherwise, for* $t \le t' = O\left(\log \frac{\phi(p^0)}{M(B)}\right)$*,* $\phi(p^t) = O\left(2^{-\Theta(t)}\phi(p^0)\right)$*, and for* $t > t'$*,* $\phi(p^t) = O\left(\frac{M(B)}{t - t'}\right)$*.*

(b) If ϕ *is strongly convex with parameter* c*,*[4] *then* $\phi(p^t) \le \left(1 - \Theta\left(\frac{c}{\overline{\gamma}}\right)\right)^t \cdot \phi(p^0)$*.*

Definition 2. *The variables* γ_j^t *and* H_{jk} *are said to be* controlled *if there are constants* $\alpha \ge 2$*,* $\epsilon_F, \epsilon_B > 0$*, with* $\frac{1}{\alpha} + 2\epsilon_B + 2\epsilon_F < 1$*, and for each* j *and time* t *at which* p_j *is updated, there are positive numbers* $\{\xi_k^t\}_{k \ne j}$*, such that:*

A1. (Local Lipschitz bound.)*Let* $S_j = \text{Span}\left\{p_j^{t-}, p_j^{t+}\right\}$*. For any* $p' \in p_{-j}^t \times S_j$*,*

$$\phi(p') - \phi(p^t) - \nabla_j \phi(p^t) \cdot (p'_j - p_j^t) \le \frac{\gamma_j^t}{\alpha}(p'_j - p_j^t)^2.$$

A2. (Upper bound on γ_j^t.) *For each* j*, there exists a finite positive number* $\overline{\gamma}_j$ *such that for all* t *at which an update to* p_j *occurs,* $\gamma_j^t \le \overline{\gamma}_j$*. We let* $\overline{\gamma} := \max_j \overline{\gamma}_j$*.*

A3. (Bound on nearby future Hessian entries.) $\sum_{k \ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]}\left(p_k^{\tau_k +}\right) \le \epsilon_F \gamma_j^t$*, where* $\sigma_k > t$ *is the time of the next update to* p_k*;*

A4. (Bound on recent past Hessian entries.) $\sum_{k \ne j}\left(\max_{i: k_i = k} \frac{1}{\xi_j^{\beta_i}}\right) \cdot H_{kj}^{[\tau_j,t]}\left(p_j^t\right) \le \epsilon_B \gamma_j^t$*, where the index* i *runs over all updates to coordinate* k *between times* τ_j *and* t*, and* β_i *is the time at which each such update occurs (this notation is defined precisely in Lemma 3).*

If the updates used fully up-to-date gradients, i.e. if $\Delta p_j = -\frac{\nabla_j \phi(p^t)}{\gamma_j^t}\Delta t_j$, rearranging Condition A1 would give the following lower bound on the progress (cf. Lemma 2 below):

$$\phi(p^{t-}) - \phi(p^{t+}) \ge \sum_j \frac{1}{\gamma_j^t}(\nabla_j \phi(p^t))^2 \Delta t_j - \frac{1}{\alpha \gamma_j^t}(\nabla_j \phi(p^t))^2 \Delta t_j^2 \ge \sum_j \left(1 - \frac{1}{\alpha}\right)\frac{(\nabla_j \phi(p^t))^2 \Delta t_j}{\gamma_j^t}.$$

[4] i.e. for any p_1, p_2 in its domain, $\phi(p_2) \ge \phi(p_1) + \nabla \phi(p_1) \cdot (p_2 - p_1) + \frac{c}{2}\|p_2 - p_1\|^2$.

The remaining conditions are present to cope with the lack of synchrony. Conditions A3 and A4 ensure that the "errors' in the gradients we use for the updates are not too large cumulatively. Basically, they will reduce the multiplier in the progress from $(1-\frac{1}{\alpha})$ to $(1-\frac{1}{\alpha}-2\epsilon_{\mathrm{F}}-2\epsilon_{\mathrm{B}})$. Recall that the lack of synchrony may result in bad updates. To hide the resulting temporary lack of progress and to show continued long-term progress, we use an amortized analysis which employs the following potential function.

$$\Phi(p^t,t,\tau)=\phi(p^t)-c_1\sum_j\int_{\tau_j}^{t}\frac{(g_j(t'))^2}{\overline{\gamma}_j}\,dt'+\sum_j\sum_i\xi_j^{\beta_i}\cdot H_{k_ij}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}\left[2-c_2(t-\beta_i)\right], \tag{2}$$

where $g_j(t'):=\nabla_j\phi\left(p^{t'}\right)$ and $\sigma_j>\tau_j$ is the time of the next update to p_j; for each j, the index i runs over all updates, between times τ_j and t, to coordinates other than j; c_1 and c_2 are positive constants whose values we will determine later. $\left\{\xi_j^{\beta_i}\right\}$ are the positive numbers in Conditions A3 and A4; note that these variables are indexed by i but not by the update coordinate k_i, so for any j, $\xi_j^{\beta_{i_1}}$ may be different from $\xi_j^{\beta_{i_2}}$, even if $k_{i_1}=k_{i_2}$.

The integral in the above potential function reflects the ideal progress were there a continuous synchronized updating of the prices, and the additional terms are present to account for the attenuation of progress due to asynchrony.

Our method of analysis is to show that $\frac{d\Phi}{dt}\le-\beta_1\Phi^2$ for a suitable constant $\beta_1>0$ whenever there is no price update, and that Φ only decreases when there is a price update; this then yields Theorem 1(a). Theorem 1(b) follows from a stronger bound on the derivative, namely that $\frac{d\Phi}{dt}\le-\beta_2\Phi$, where $\beta_2>0$. This general approach for asynchrony analysis was used previously by Cheung et al. [7] for a result in the style of (b), but for a quite different potential function.

It is straightforward to show that when there is no update,

$$\frac{d\Phi}{dt}=-c_1\sum_j\frac{(g_j(t))^2}{\overline{\gamma}_j}-c_2\sum_j\sum_i\xi_j^{\beta_i}\cdot H_{k_ij}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. \tag{3}$$

Lemma 2 below bounds the change to ϕ when there is an update. Lemma 3 states some useful bounds on the maximum change that can occur to the gradient between two updates to the same coordinate. Lemma 4 below bounds the change to Φ when there is an update.

Lemma 2. *Suppose there is an update to p_j at time t according to rule* (1)*, with γ_j^t satisfying Condition A1. Let ϕ^- and ϕ^+ denote, respectively, the convex function values just before and just after the update. Let $g_j:=\nabla_j\phi(p^t)$ and $\tilde{g}_j\equiv\tilde{g}_j(t)$. Let Δp_j be the change to p_j made by the update, i.e. $\Delta p_j:=-\frac{\tilde{g}_j(t)}{\gamma_j^t}\Delta t_j$. Then*

$$\phi^--\phi^+\ge\left(1-\frac{1}{\alpha}\right)\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j}-|g_j-\tilde{g}_j|\cdot|\Delta p_j|.$$

Lemma 3. *Suppose that between times τ_j and t, there are updates to the sequence of coordinates $k_1,k_2,\cdots,k_m$, which may include repetitions, but include no update to coordinate j. Let $\beta_1,\beta_2,\cdots,\beta_m$ denote the times at which these updates occur. Let $\tilde{g}_{j,\max}$ and $\tilde{g}_{j,\min}$ denote,* *respectively, the maximum and minimum values of $\nabla_j(p')$, where $p' \in \tilde{P}_j^{[\tau_j,t]}\left(p_j^t\right)$. For any positive numbers $\{\eta_i\}_{i=1\cdots m}$, for each $k \neq j$, let $\bar{\eta}_k := \min_{i:k_i=k} \eta_i$. Then for any real number μ,*

$$|\mu| \cdot (\tilde{g}_{j,\max} - \tilde{g}_{j,\min}) \leq 2\mu^2 \sum_{k\neq j} \frac{1}{\bar{\eta}_k} H_{kj}^{[\tau_j,t]}\left(p_j^t\right) + \sum_{i=1}^{m} \eta_i \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \tag{4}$$

and

$$(\tilde{g}_{j,\max} - \tilde{g}_{j,\min})^2 \leq 8\left(\sum_{i=1}^{m} \eta_i \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}\right)\left(\sum_{k\neq j} \frac{1}{\bar{\eta}_k} H_{kj}^{[\tau_j,t]}\left(p_j^t\right)\right). \tag{5}$$

Lemma 4. *Suppose that there is an update to p_j at time t. Suppose that γ_j^t is chosen so that Conditions A1, A3 and A4 hold. Let Φ^- and Φ^+, respectively, denote the values of Φ just before and just after the update. Then*

$$\Phi^- - \Phi^+ \geq \left(1 - \frac{1}{\alpha} - 2\epsilon_B - c_1(1+4\epsilon_B) - 2\epsilon_F\right) \frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} + (1 - c_2 - c_1(2+8\epsilon_B)) \sum_{i=1}^{m} \xi_j^{\beta_i} \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

Proof: By Lemma 2 and the fact $(t-\beta_i) \leq (t-\tau_j) \leq 1$,

$$\begin{aligned}
\Phi^- - \Phi^+ &= \phi^- - \phi^+ - c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j}\,dt' + \sum_i \xi_j^{\beta_i} \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \left[2 - c_2(t-\beta_i)\right] \\
&\qquad - 2\sum_{k\neq j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]}\left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j} \\
&\geq \left(1 - \frac{1}{\alpha}\right) \frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - \underbrace{|g_j - \tilde{g}_j| \cdot |\Delta p_j|}_{E_1} - \underbrace{c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j}\,dt'}_{E_2} \\
&\qquad + (2-c_2)\sum_i \xi_j^{\beta_i} \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} - \underbrace{2\sum_{k\neq j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]}\left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j}}_{E_3}.
\end{aligned} \tag{6}$$

We bound E_1, E_2 and E_3 below. We will be applying (4) and (5) with $\eta_i = \xi_j^{\beta_i}$. Let

$$V_1 := \sum_{k\neq j} \frac{1}{\min_{i:k_i=k} \xi_j^{\beta_i}} H_{kj}^{[\tau_j,t]}\left(p_j^t\right) \qquad \text{and} \qquad V_2 := \sum_{i=1}^{m} \xi_j^{\beta_i} \cdot H_{k_ij}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

Note that by Condition A4, $V_1 \leq \epsilon_B \gamma_j^t$. By (4), $E_1 \leq 2(\Delta p_j)^2 V_1 + V_2 \leq 2\epsilon_B \gamma_j^t (\Delta p_j)^2 + V_2$.

To bound E_2, first note that for any $t' \in (\tau_j, t]$, $p^{t'} \in \tilde{P}_j^{[\tau_j,t]}\left(p_j^t\right)$. Then

$$
\begin{aligned}
&\frac{(g_j(t'))^2}{\overline{\gamma}_j} - \frac{(\tilde{g}_j)^2}{\overline{\gamma}_j} = \frac{(g_j(t') - \tilde{g}_j)^2}{\overline{\gamma}_j} - \frac{2\tilde{g}_j}{\overline{\gamma}_j}(\tilde{g}_j - g_j(t')) \\
&\leq \frac{(g_j(t') - \tilde{g}_j)^2}{\overline{\gamma}_j} + 2\left|\frac{\tilde{g}_j}{\overline{\gamma}_j}\right| \cdot |\tilde{g}_j - g_j(t')| \;\leq\; \frac{8}{\overline{\gamma}_j} V_2 V_1 + \frac{4(\tilde{g}_j)^2}{(\overline{\gamma}_j)^2} V_1 + 2V_2 \qquad \text{(by Eqns. (5) and (4))} \\
&\leq \frac{8\epsilon_{\mathrm{B}}\gamma_j^t}{\overline{\gamma}_j} V_2 + \frac{4\epsilon_{\mathrm{B}}\gamma_j^t(\tilde{g}_j)^2}{(\overline{\gamma}_j)^2} + 2V_2 \;\leq\; \frac{4\epsilon_{\mathrm{B}}(\tilde{g}_j)^2}{\overline{\gamma}_j} + (2 + 8\epsilon_{\mathrm{B}})V_2 \qquad \text{(by Condition A2)} \qquad (7)
\end{aligned}
$$

Hence $\frac{(g_j(t'))^2}{\overline{\gamma}_j} \leq (1 + 4\epsilon_{\mathrm{B}})\frac{(\tilde{g}_j)^2}{\overline{\gamma}_j} + (2 + 8\epsilon_{\mathrm{B}})V_2$, and then as $\Delta t_j \leq 1$,

$$
E_2 \;\leq\; c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j} dt' \;\leq\; c_1(1 + 4\epsilon_{\mathrm{B}})\frac{(\tilde{g}_j)^2 \Delta t_j}{\overline{\gamma}_j} + c_1(2 + 8\epsilon_{\mathrm{B}})V_2 \tag{8}
$$

Finally, by Condition A3, $E_3 \leq 2\epsilon_{\mathrm{F}}\gamma_j^t \frac{(\Delta p_j)^2}{\Delta t_j}$.

Combining the above bounds on E_1, E_2, E_3 yields

$$
\begin{aligned}
\Phi^- - \Phi^+ \geq &\left(1 - \frac{1}{\alpha}\right)\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - \left[2\epsilon_{\mathrm{B}}\gamma_j^t(\Delta p_j)^2 + V_2\right] - \left[c_1(1 + 4\epsilon_{\mathrm{B}})\frac{(\tilde{g}_j)^2 \Delta t_j}{\overline{\gamma}_j} + c_1(2 + 8\epsilon_{\mathrm{B}})V_2\right] \\
&+ (2 - c_2)V_2 - 2\epsilon_{\mathrm{F}}\gamma_j^t \frac{(\Delta p_j)^2}{\Delta t_j}.
\end{aligned}
$$

As $\Delta p_j = -\frac{\tilde{g}_j(t)}{\gamma_j^t}\Delta t_j$ and $\Delta t_j \leq 1$, the result follows. $\square$

Lemma 5. *If* $2 - c_2 \geq c_1(2 + 8\epsilon_B)$, *then* $\Phi(p^t, t, \tau) \geq [1 - 2c_1(1 + 4\epsilon_B)]\,\phi(p^t)$.

Proof of Theorem 1(a): Choose $c_1 = (1 + 4\epsilon_{\mathrm{B}})^{-1} \cdot \min\left\{1 - \frac{1}{\alpha} - 2\epsilon_{\mathrm{B}} - 2\epsilon_{\mathrm{F}}, \frac{1}{4}\right\}$ and $c_2 = 1 - c_1(2 + 8\epsilon_{\mathrm{B}})$. Then the following hold: (i) $c_1, c_2 > 0$; (ii) $1 - \frac{1}{\alpha} - 2\epsilon_{\mathrm{B}} - 2\epsilon_{\mathrm{F}} - c_1(1 + 4\epsilon_{\mathrm{B}}) \geq 0$; (iii) $1 - c_2 - c_1(2 + 8\epsilon_{\mathrm{B}}) = 0$; (iv) $2 - c_2 \geq c_1(2 + 8\epsilon_{\mathrm{B}})$; (v) $c_1(1 + 4\epsilon_{\mathrm{B}}) \leq \frac{1}{4}$.

By (ii), (iii) and Lemma 4, Φ does not increase at any update.

By (iv), (v) and Lemma 5, $\Phi(p^t, t, \tau) \geq \frac{\phi(p^t)}{2}$. Thus, $\forall t \geq 0$, $\phi(p^t) \leq 2\Phi(p^t, t, \tau) \leq 2\Phi(p^0, 0, \vec{0}) = 2\phi(p^0)$, i.e. $\{p^t\}_{t \geq 0}$ is contained in the set $\{p' \mid \phi(p') \leq 2\phi(p^0)\}$, which, by assumption, has diameter at most B.

Note that at any time t, by the convexity of ϕ, $\phi(p^t) + \sum_j g_j(t) \cdot (p_j^* - p_j^t) \leq \phi^* = 0$ and hence

$$
\sum_j |g_j(t)| \cdot |p_j^t - p_j^*| \geq \sum_j g_j(t) \cdot (p_j^t - p_j^*) \geq \phi(p^t) \geq 0.
$$

By the Cauchy-Schwarz inequality,

$$
\phi(p^t) \leq \sum_j |g_j(t)| \cdot |p_j^t - p_j^*| \leq \sqrt{\left(\sum_j (g_j(t))^2\right)\left(\sum_j (p_j^t - p_j^*)^2\right)} \leq B\sqrt{\sum_j (g_j(t))^2}.
$$

Then

$$\sum_j \frac{(g_j(t))^2}{\overline{\gamma}_j} \geq \frac{1}{\overline{\gamma}} \sum_j (g_j(t))^2 \geq \frac{1}{\overline{\gamma}} \left(\frac{\phi(p^t)}{B} \right)^2 = \frac{1}{B^2 \overline{\gamma}} \phi(p^t)^2.$$

By (3),

$$\frac{d\Phi}{dt} \leq -\frac{c_1}{B^2 \overline{\gamma}} \cdot \phi(p^t)^2 - c_2 \sum_j \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +} \right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

By (2), $\Phi(p^t, t, \tau) \leq \phi(p^t) + 2 \sum_j \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +} \right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}$. Let $X_1 := \phi(p^t)$ and $X_2 := \sum_j \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +} \right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}$. Then $\Phi \leq X_1 + 2X_2$ and $\frac{d\Phi}{dt} \leq -\frac{c_1}{B^2 \overline{\gamma}} (X_1)^2 - c_2 X_2$. Let $M(B) := \Theta(B^2 \overline{\gamma})$. As $\phi(p^t) \leq 2\Phi(t)$, this guarantees that if $\phi(p^0) = \Phi(p^0) \leq M(B)$, then $\phi(p^t) = O\left(\frac{M(B)}{t}\right)$; and otherwise, for $t \leq t' = O\left(\log \frac{\phi(p^0)}{M(B)}\right)$, $\phi(p^t) = O\left(2^{-\Theta(t)} \phi(p^0)\right)$, and for $t > t'$, $\phi(p^t) = O\left(\frac{M(B)}{t - t'}\right)$. □

Proof of Theorem 1(b): If ϕ is strongly convex with parameter c, then, by definition,

$$\begin{aligned} 0 = \phi^* &\geq \phi(p^t) + \sum_j g_j(t) \cdot (p_j^* - p_j^t) + \frac{c}{2} \sum_j (p_j^* - p_j^t)^2 \\ &\geq \phi(p^t) + \min_{p'} \left\{ \sum_j g_j(t) \cdot (p_j' - p_j^t) + \frac{c}{2} (p_j' - p_j^t)^2 \right\}. \end{aligned}$$

Computing the minimum point of the quadratic polynomial in $(p_j' - p_j^t)$ yields $0 \geq \phi(p^t) - \sum_j \frac{(g_j(t))^2}{2c}$. Then

$$\sum_j \frac{(g_j(t))^2}{\overline{\gamma}_j} \geq \frac{1}{\overline{\gamma}} \sum_j (g_j(t))^2 \geq \frac{2c}{\overline{\gamma}} \phi(p^t).$$

As in Case (a), $\Phi \leq X_1 + 2X_2$; and by (3), $\frac{d\Phi}{dt} \leq -\frac{2cc_1}{\overline{\gamma}} X_1 - c_2 X_2$. This guarantees that $2\phi(p^t) \leq \Phi(t) \leq (1 - \delta(c))^t \phi(p^0)$, where $\delta(c) = \min\{\frac{cc_1}{\overline{\gamma}}, \frac{c_2}{4}\}$. □

4 Solving Matrix Systems

For any symmetric and positive definite (SPD) matrix $A \in \mathbb{R}^{n \times n}$ and $b, p \in \mathbb{R}^n$, let $f_{A,b}(p) = \frac{1}{2} p^{\mathrm{T}} A p - p^{\mathrm{T}} b$. It is well known that $f_{A,b}(p)$ is a strictly convex function of p, and $\nabla f_{A,b}(p) = Ap - b$. Therefore, finding the minimum point of $f_{A,b}(p)$ is equivalent to solving the linear system $Ap = b$, and hence one can solve the linear system by performing gradient descent on $f_{A,b}(p)$.

The Hessian of $f_{A,b}(p)$ is $\nabla^2 f_{A,b}(p) = A$, a constant matrix. This allows a simple rule to determine a *constant* step size for each coordinate. By taking all the ξ values to be 1, to apply Theorem 1, it suffices to have $\gamma_j^t = \gamma_j$ satisfy $\gamma_j \geq \frac{A_{jj}}{2} \alpha$ (for A1), $\frac{4}{\gamma_j} \sum_{k \neq j} |A_{jk}| < 1 - \frac{1}{\alpha}$ (combining A3, A4 and the bound $\frac{1}{\alpha} + 2\epsilon_{\mathrm{F}} + 2\epsilon_{\mathrm{B}} < 1$), and $\alpha \geq 2$. These imply it suffices that the step size, $1/\gamma_j$, be less than $\left[\max \left\{ \frac{A_{jj} + 8 \sum_{k \neq j} |A_{kj}|}{2}, A_{jj} \right\} \right]^{-1}$.

Another application is given by the following class of optimization problems (see Nesterov [14]): minimizing $F(p) := \sum_{i=1}^{n} f_i(p_i) + \frac{1}{2}\|Ap-b\|^2$, where the f_i are convex differentiable univariate functions, $A \in \mathbb{R}^{r\times n}$ is an $r\times n$ real matrix and $b \in \mathbb{R}^r$. The Hessian of F at p is $A^{\mathrm{T}}A+D$, where D is the diagonal matrix with $D_{jj} = f_j''(p_j)$. If $f_j''(p)$ is bounded by L_j, again, it suffices to have $\gamma_j^t = \gamma_j$ satisfy $\gamma_j \geq \frac{(A^{\mathrm{T}}A)_{jj}+L_j}{2}\alpha$, $\frac{4}{\gamma_j}\sum_{k\neq j}|(A^{\mathrm{T}}A)_{jk}| < 1-\frac{1}{\alpha}$, and $\alpha \geq 2$. These imply it suffices that the step size, $1/\gamma_j$, be less than $\left[\max\left\{\frac{(A^{\mathrm{T}}A)_{jj}+L_j+8\sum_{k\neq j}|(A^{\mathrm{T}}A)_{kj}|}{2}, (A^{\mathrm{T}}A)_{jj}+L_j\right\}\right]^{-1}$.

Next, we discuss how $\nabla_j F(p)$ is computed by the j-th core. Let $G(p) = Ap - b$ and let A_j denote the j-th column of the matrix A. Then $\nabla_j F(p) = f_j'(p_j) + (A_j)^{\mathrm{T}}G(p)$. $f_j'(p_j)$ is recomputed only when p_j changes. For any k, when p_k is changed by Δp_k, $G(p+\Delta p_k)-G(p) = \Delta p_k A_k$, and hence $(A_j)^{\mathrm{T}}G(p)$ changes by $\Delta p_k (A_j)^{\mathrm{T}}A_k$. Note that $(A_j)^{\mathrm{T}}A_k$ is a constant and hence can be pre-calculated, so the above equation provides a quick way to update $\nabla_j F(p)$ once the j-th core receives the message with Δp_k.

Recall that our AGD algorithm allows different coordinate values to be variously dated, under the constraint that they are all no older than the time of the last update. It is natural to aim to have essentially the same frequency of update for each coordinate. Accordingly, at the i-th round of updates, each core can simply ensure it has received the update for the previous round from every other core. The update messages might arrive at different times, but the j-th core needs not wait until it collects all such messages. It can simply compute the changes to $\nabla_j F(p)$ incrementally as it receives updates Δp_k to p_k. This avoids the need for any explicit synchronization.

5 Tatonnement in Fisher Markets

A *Fisher market* comprises a set of n goods and two sets of agents, sellers and buyers. The sellers bring the goods to market and the buyers bring money with which to buy the goods. The trade is driven by a collection of non-negative prices $\{p_j\}_{j=1\cdots n}$, one price per good. WLOG, we assume that each seller brings one distinct good to the market, and she is the price-setter for this good. By normalization, we may assume that each seller brings one unit of her good to the market.

Each buyer i starts with e_i money, and has a utility function $u_i(x_{i1}, x_{i2}, \cdots, x_{in})$ expressing her preferences: if she prefers bundle $\{x_{ij}^a\}_{j=1\cdots n}$ to bundle $\{x_{ij}^b\}_{j=1\cdots n}$, then $u_i(\{x_{ij}^a\}_{j=1\cdots n}) > u_i(\{x_{ij}^b\}_{j=1\cdots n})$. At any given prices $\{p_j\}_{j=1\cdots n}$, each buyer i seeks to purchase a maximum utility bundle of goods costing at most e_i. The *demand* for good j, denoted by x_j, is the total quantity of the good sought by all buyers. The *supply* of good j is the quantity of good j its seller brings to the market, which we have assumed to be 1. The *excess demand* for good j, denoted by z_j, is the demand for the good minus its supply, i.e. $z_j = x_j - 1$. Prices $\{p_j^*\}_{j=1\cdots n}$ are said to form a *market equilibrium* if, for any good j with $p_j^* > 0$, $z_j = 0$, and for any good j with $p_j^* = 0$, $z_j \leq 0$.

The following two classes of utility functions are commonly used in market models. The first class is the Constant Elasticity of Substitution (CES) utility function:

$$u_i\left(x_{i1}, x_{i2}, \cdots, x_{in}\right) = \left(a_{i1}(x_{i1})^{\rho_i} + a_{i2}(x_{i2})^{\rho_i} + \cdots + a_{in}(x_{in})^{\rho_i}\right)^{1/\rho_i},$$

where $\rho_i \leq 1$ and $\forall j$, $a_{ij} \geq 0$. $\theta_i := \rho_i/(\rho_i - 1)$ is a parameter which will be used in the analysis. In this paper we focus on the cases $\rho_i \leq 0$, in which goods are complements and hence the utility function is called a complementary-CES utility function. It is easy to extend our analysis to the cases $\rho_i \geq 0$, which had been analysed in [10, 11]. The second class is the Leontief utility function:

$$u_i\left(x_{i1}, x_{i2}, \cdots, x_{in}\right) = \min_{j \in S} \left\{b_{ij} x_{ij}\right\},$$

where S is a non-empty subset of the goods in the market, and $\forall j \in S$, $b_{ij} > 0$.

Cheung, Cole and Devanur [6] showed that tatonnement is equivalent to gradient descent on a convex function ϕ for Fisher markets with buyers having complementary-CES or Leontief utility functions (defined in the appendix). To be specific, $\nabla_j \phi(p) = -z_j(p)$, and the convex function ϕ is $\phi(p) = \sum_j p_j + \sum_i \hat{u}_i(p)$, where $\hat{u}_i(p)$ is the optimal utility that buyer i can attain at prices p. The corresponding update rule is

$$p'_j = p_j \cdot \left(1 + \lambda \cdot \min\{\tilde{z}_j, 1\} \cdot (t - \tau_j)\right), \tag{9}$$

where $\tilde{z}_j$ is a value between the minimum and maximum excess demands during the time interval $(\tau_j, t]$, and $\lambda > 0$ is a suitable constant. As the update rule is multiplicative, we assume that the initial prices are positive.

Note that $\gamma_j^t = \frac{\max\{1, \tilde{z}_j\}}{\lambda p_j}$. As we will see, it suffices that $\lambda \leq \frac{1}{23.46}$. In comparison, in the synchronous version, $\gamma_j^t \geq \frac{6 \max\{1, z_j^t\}}{p_j}$, so the step sizes of the asynchronous tatonnement are a constant fraction of those used in its synchronous counterpart.

Theorem 6. *For $\lambda \leq \frac{1}{23.46}$, asynchronous tatonnement price updates using rule* (9) *converge toward the market equilibrium in any complementary-CES or Leontief Fisher market.*

In a Fisher market with buyers having complementary-CES utility functions, Properties 1 and 2 below are well-known. Property 3 was proved in [6] and implies that Condition A1 holds when $\alpha = 6$ and $\gamma_j^t \geq 9.5 x_j(p^t)/p_j^t$.

1. Let $x_{i\ell}(p)$ denote the buyer i's demand for good ℓ at prices p. Then for $k \neq j$,

$$\left|\frac{\partial^2 \phi}{\partial p_j \partial p_k}\right| = \sum_i \frac{\theta_i x_{ij}(p) x_{ik}(p)}{e_i} \leq \sum_i \frac{x_{ij}(p) x_{ik}(p)}{e_i}.$$

2. Given positive prices p, for any $0 < r_1 < r_2$, let p' be prices such that for all j, $r_1 p_j \leq p'_j \leq r_2 p_j$. Then for all j, $\frac{1}{r_2} x_j(p) \leq x_j(p') \leq \frac{1}{r_1} x_j(p)$.

3. If $\frac{\Delta p_j}{p_j} \leq 1/6$, then $\phi(p + \Delta p) - \phi(p) - \nabla_j \phi(p) \cdot \Delta p_j \leq \frac{1.5 x_j}{p_j} (\Delta p_j)^2$.

We outline the analysis for the complementary-CES case. As $\lambda \leq \frac{1}{23.46}$, within one unit of time, each price can vary by a factor between $(9/10)^2 = 81/100$ and $(11/10)^2 = 121/100$.[5] Hence, within one unit of time, the demand can vary by a factor between 100/121 and 100/81.

For each update to p_j at time t, we choose $\xi_k^t := p_k^t / p_j^t$. Then the following lemma bounds the sums in Conditions A3 and A4.

[5]These bounds are loose, but they suffice for our purpose.

Lemma 7. (a) $\sum_{k\neq j} \xi_k^t \cdot H_{jk}^{[t,\sigma_j]}\left(p_k^{\tau_k+}\right) \leq \frac{1.53 x_j(p^t)}{p_j^t}$;
(b) $\sum_{k\neq j}\left(\max_{q:k_q=k} \frac{1}{\xi_j^{\beta_q}}\right) \cdot H_{kj}^{[\tau_j,t]}\left(p_j^t\right) \leq \frac{1.89 x_j(p^t)}{p_j^t}$.

Proof:

$$
\begin{aligned}
\sum_{k\neq j} \xi_k^t \cdot H_{jk}^{[t,t+1]}\left(p_k^{\tau_k+}\right) &= \sum_{k\neq j} \frac{p_k^t}{p_j^t} \cdot \max_{p'\in \tilde{P}_k^{[t,t+1]}\left(p_k^{\tau_k+}\right)} \left|\frac{\partial^2\phi}{\partial p_j \partial p_k}\right| \\
&\leq \frac{1}{p_j^t} \sum_{k\neq j} p_k^t \cdot \max_{p'\in \tilde{P}_j^{[t,t+1]}\left(p_k^{\tau_k+}\right)} \sum_i \frac{x_{ij}(p')x_{ik}(p')}{e_i} \leq \frac{1}{p_j^t}\sum_{k\neq j} p_k^t \sum_i \frac{\left(\frac{100}{81}x_{ij}(p^t)\right)\left(\frac{100}{81}x_{ik}(p^t)\right)}{e_i} \\
&\leq \frac{1.53}{p_j^t}\sum_i x_{ij}(p^t)\sum_{k\neq j}\frac{p_k^t x_{ik}(p^t)}{e_i} \leq \frac{1.53}{p_j^t}\sum_i x_{ij}(p^t) = \frac{1.53 x_j(p^t)}{p_j^t}.
\end{aligned}
$$

And

$$
\begin{aligned}
\sum_{k\neq j}\left(\max_{q:k_q=k}\frac{1}{\xi_j^{\beta_q}}\right) \cdot H_{kj}^{[\tau_j,t]}\left(p_j^t\right) &= \sum_{k\neq j}\frac{\max_{q:k_q=k} p_k^{\beta_q}}{p_j^t} \cdot \max_{p'\in \tilde{P}_j^{[\tau_j,t]}\left(p_j^t\right)} \left|\frac{\partial^2\phi}{\partial p_j \partial p_k}\right| \\
&\leq \frac{1}{p_j^t}\sum_{k\neq j}\left(\frac{100}{81}p_k^t\right)\sum_i \frac{\left(\frac{100}{81}x_{ij}(p^t)\right)\left(\frac{100}{81}x_{ik}(p^t)\right)}{e_i} \\
&\leq \frac{1.89}{p_j^t}\sum_i x_{ij}(p^t)\sum_{k\neq j}\frac{p_k^t x_{ik}(p^t)}{e_i} \leq \frac{1.89}{p_j^t}\sum_i x_{ij}(p^t) = \frac{1.89 x_j(p^t)}{p_j^t}.
\end{aligned}
$$

$\square$

Proof of Theorem 6 for the CES case: By Property 3, Condition A1 is satisfied by setting $\gamma_j^t \geq \frac{9.5 x_j(p^t)}{p_j^t}$ and $\alpha = 6$. By Lemma 7, Conditions A3 and A4 are satisfied by setting $\epsilon_{\mathrm{F}} = 1/6$ and $\epsilon_{\mathrm{B}} = 1/5$, and $1 - \frac{1}{\alpha} - 2\epsilon_{\mathrm{F}} - 2\epsilon_{\mathrm{B}} = \frac{1}{10} > 0$.

As discussed in [10], the seller might know only $\tilde{x}_j$ but not x_j. As $\tilde{x}_j \geq \frac{81}{100}x_j$, it would be more natural to use $\gamma_j^t \geq \frac{11.73\tilde{x}_j}{p_j}$, or the even weaker (but still more natural) $\gamma_j^t \geq \frac{23.46\max\{1,\tilde{z}_j\}}{p_j}$, which yields update rule (9).

[6] proved that prices in tatonnement cannot get arbitrarily close to zero and hence demands cannot increase indefinitely, so $\overline{\gamma}_j$, as defined in Condition A2, is finite. [6] also showed that ϕ is strongly convex. The result follows from Theorem 1(b). $\square$

Ongoing Complementary-CES Fisher Markets Cole and Fleischer's Ongoing market model [10] incorporates asynchronous tatonnement and warehouses to form a self-contained dynamic market model. The price update rule is designed to achieve two goals simultaneously: convergence toward the market equilibrium and warehouse "balance". As in [7], we modify the price update rule (9) to achieve both targets. Analysing its convergence entails the design of a significantly more involved potential function; the details are given in the appendix.

Leontief Fisher Markets It is well-known that Leontief utility functions can be considered as the "limit" of CES utility functions as $\rho \to -\infty$. Our analysis for CES Fisher markets can be reused, with no modification needed, to show that in any Leontief Fisher market, $\Phi(p^t, t, \tau)$ decreases with t. However, as an equilibrium price in a Leontief Fisher market can be zero, it is unavoidable that the chosen step size γ_j^t may tend to infinity (as $\gamma_j^t = \Omega(1/p_j)$), violating Condition A2; thus Theorem 1 cannot be applied directly.

On top of the result that $\Phi(p^t, t, \tau)$ decreases with t, we provide additional arguments to show that tatonnement with update rule (9) still converges toward the market equilibrium in Leontief Fisher markets. The proof is given in the appendix. However, this result does not provide a bound on the rate of convergence, which appears to preclude incorporating warehouses into the analysis.

Further Discussion of Asynchronous Dynamics Computer science has long been concerned with the organization and manipulation of information in the form of well-defined problems with a clear intended outcome. But in the last 15 years, computer science has gained a new dimension, in which outcomes are predicted or described, rather than designed. Examples include bird flocking [4], influence systems [5], spread of information memes across the Internet [13] and market economies [10]. Many of these problems fall into the broad category of analysing dynamic systems. Dynamic systems are a staple of the physical sciences; often the dynamics are captured via a neat, deterministic set of rules (e.g. Newton's law of motion, Maxwell's equations for electrodynamics). The modeling of dynamic systems with intelligent agents presents new challenges because agent behavior may not be wholly consistent or systematic. One issue that has received little attention is the timing of agents' actions. Typically, a fixed schedule has been assumed (e.g. synchronous or round robin), perhaps because it was more readily analysed.

This work provides a second demonstration (the first demonstration is in [11, 7]) and further development of a method for analysing asynchronous dynamics, here for dynamics which are equivalent to gradient descent. This methodology may be of wider interest.

References

[1] Kenneth J. Arrow, H. D. Block, and Leonid Hurwicz. On the stability of competitive equilibrium, ii. *Econometrica*, 27(1):82–109, 1959.

[2] Dimitri P. Bertsekas and John N. Tsitsiklis. Gradient convergence in gradient methods with errors. *SIAM J. Optimization*, 10(3):627–642, 2000.

[3] Vivek S. Borkar. Asynchronous stochastic approximations. *SIAM J. Control and Optimization*, 36(3):662–663, 1998.

[4] Bernard Chazelle. Natural algorithms. In *SODA*, pages 422–431, 2009.

[5] Bernard Chazelle. The dynamics of influence systems. In *FOCS*, pages 311–320, 2012.

[6] Yun Kuen Cheung, Richard Cole, and Nikhil R. Devanur. Tatonnement beyond gross substitutes? gradient descent to the rescue. In *STOC*, pages 191–200, 2013.

[7] Yun Kuen Cheung, Richard Cole, and Ashish Rastogi. Tatonnement in ongoing markets of complementary goods. In *EC*, pages 337–354, 2012.

[8] Bruno Codenotti, Benton McCune, and Kasturi R. Varadarajan. Market equilibrium via the excess demand function. In *STOC*, pages 74–83, 2005.

[9] Bruno Codenotti and Kasturi R. Varadarajan. Efficient computation of equilibrium prices for markets with leontief utilities. In *ICALP*, pages 371–382, 2004.

[10] Richard Cole and Lisa Fleischer. Fast-converging tatonnement algorithms for one-time and ongoing market problems. In *STOC*, pages 315–324, 2008.

[11] Richard Cole, Lisa Fleischer, and Ashish Rastogi. Discrete price updates yield fast convergence in ongoing markets with finite warehouses. *CoRR*, 2010.

[12] John Langford, Alex J. Smola, and Martin Zinkevich. Slow learners are fast. In *NIPS*, pages 2331–2339, 2009.

[13] Jure Leskovec, Lars Backstrom, and Jon M. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *KDD*, pages 497–506, 2009.

[14] Yu. Nesterov. Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM J. Optimization*, 22(2):341–362, 2012.

[15] Feng Niu, Benjamin Recht, Christopher Re, and Stephen J. Wright. Hogwild: A lock-free approach to parallelizing stochastic gradient descent. In *NIPS*, pages 693–701, 2011.

[16] Christos H. Papadimitriou and Mihalis Yannakakis. An impossibility theorem for price-adjustment mechanisms. *PNAS*, 5(107):1854–1859, 2010.

[17] John N. Tsitsiklis, Dimitri P. Bertsekas, and Michael Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.

[18] Hirofumi Uzawa. Walras' tatonnement in the theory of exchange. *Review of Economic Studies*, 27(3):182–194, 1960.

[19] Léon Walras. *Eléments d'Economie Politique Pure*. Corbaz, 1874. (Translated as: *Elements of Pure Economics.* Homewood, IL: Irwin, 1954.).

A Missing Proofs in Section 3

Proof of Lemma 2: By Condition A1, $\phi^+ - \phi^- - g_j \Delta p_j \leq \frac{\gamma_j^t}{\alpha}(\Delta p_j)^2$. Then

$$\begin{aligned}
\phi^- - \phi^+ &\geq -[\tilde{g}_j + (g_j - \tilde{g}_j)]\Delta p_j - \frac{\gamma_j^t}{\alpha}(\Delta p_j)^2 \\
&\geq \frac{\gamma_j^t \Delta p_j}{\Delta t_j} \cdot \Delta p_j - \frac{1}{\alpha} \cdot \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} - |g_j - \tilde{g}_j| \cdot |\Delta p_j| \qquad (\text{as } \Delta t_j \leq 1) \\
&= \left(1 - \frac{1}{\alpha}\right) \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} - |g_j - \tilde{g}_j| \cdot |\Delta p_j|.
\end{aligned}$$

□

Proof of Lemma 3: We begin by showing

$$\tilde{g}_{j,\max} - \tilde{g}_{j,\min} \leq 2 \sum_{i=1}^{m} H_{k_i j}^{[\beta_i, t]}\left(p_j^t\right) \cdot |\Delta p_{k_i}|. \tag{10}$$

First of all, we define a few useful notations. Let $\tilde{p}_{\max}$ and $\tilde{p}_{\min}$, respectively, denote the $\tilde{p}$-values at which $\nabla_j \phi(\tilde{p})$ yields $\tilde{g}_{j,\max}$ and $\tilde{g}_{j,\min}$. Let $p_{k,\min}^{(t_1,t]} := \min_{t' \in (t_1, t]} p_k^{t'}$ and $p_{k,\max}^{(t_1,t]} := \max_{t' \in (t_1,t]} p_k^{t'}$. Let $\beta_0 := \tau_j$.

To prove (10), we first construct a path P that connects $\tilde{p}_{\max}$ and $\tilde{p}_{\min}$, with each edge in P corresponding to a price update between times τ_j and t. The construction builds two paths, P^s, starting at $\tilde{p}_{\max}$, and P^e, starting at $\tilde{p}_{\min}$. Note that $\tilde{p}_{\max}, \tilde{p}_{\min} \in \tilde{P}_j^{[\tau_j, t]}\left(p_j^t\right)$, and for all $k \neq j$, $(\tilde{p}_{\max})_k, (\tilde{p}_{\min})_k \in \left[p_{k,\min}^{(\beta_0,t]}, p_{k,\max}^{(\beta_0,t]}\right]$. P^s and P^e will be constructed in m steps that correspond to the m price updates at times $\beta_1, \beta_2, \cdots, \beta_m$. By the end of the ℓ-th step, our construction ensures that the end points of P^s and P^e are in the set $\tilde{P}_j^{[\beta_\ell, t]}\left(p_j^t\right)$. Hence, by the end of the m-th step, the end points of P^s and P^e are in the set $\tilde{P}_j^{[\beta_m, t]}\left(p_j^t\right)$, which is a singleton, so the two end points must be equal. This allows P^s and P^e to be concatenated at their end points to form the path P. The specifics of the construction are as follows:

1. Let $\mathring{p}^s$ and $\mathring{p}^e$, respectively, denote the end points of P^s and P^e, i.e. initially, $\mathring{p}^s = \tilde{p}_{\max}$ and $\mathring{p}^e = \tilde{p}_{\min}$.

2. For $i = 1 \cdots m$, do:

   - Suppose span $\left\{\mathring{p}_{k_i}^s, \mathring{p}_{k_i}^e\right\} = [l_i, r_i]$. WLOG, suppose that $\mathring{p}_{k_i}^s = l_i$.[6]

     Note that by the end of the last step, the construction ensures that $l_i, r_i \in \left[p_{k_i,\min}^{(\beta_{i-1},t]}, p_{k_i,\max}^{(\beta_{i-1},t]}\right]$.

     Also, note that at most one of the strict inequalities $p_{k_i,\min}^{(\beta_i,t]} > p_{k_i,\min}^{(\beta_{i-1},t]}$ and $p_{k_i,\max}^{(\beta_i,t]} < p_{k_i,\max}^{(\beta_{i-1},t]}$ holds, and hence $l_i < p_{k_i,\min}^{(\beta_i,t]} < p_{k_i,\max}^{(\beta_i,t]} < r_i$ is not possible.

   - For any p, let $p' = (p_{-k}, x)$ be the vector such that $p'_k = x$, and for all $h \neq k$, $p'_h = p_h$. Depending on the values of $l_i, r_i, p_{k_i,\min}^{(\beta_i,t]}, p_{k_i,\max}^{(\beta_i,t]}$, there are five cases.

[6] If $\mathring{p}_{k_i}^e = l_i$, swap the roles of P^s and P^e in the current for loop.

(a) If $p^{(\beta_i,t]}_{k_i,\min} \le l_i \le r_i \le p^{(\beta_i,t]}_{k_i,\max}$, do nothing.

(b) If $l_i < p^{(\beta_i,t]}_{k_i,\min} \le r_i \le p^{(\beta_i,t]}_{k_i,\max}$, let $\mathring{p}' = \left(\mathring{p}^s_{-k_i}, p^{(\beta_i,t]}_{k_i,\min}\right)$; in P^s, connect $\mathring{p}^s$ to $\mathring{p}'$, and update $\mathring{p}^s$ to $\mathring{p}'$.

(c) If $l_i \le r_i < p^{(\beta_i,t]}_{k_i,\min} \le p^{(\beta_i,t]}_{k_i,\max}$,
- let $\mathring{p}' = \left(\mathring{p}^s_{-k_i}, p^{(\beta_i,t]}_{k_i,\min}\right)$; in P^s, connect $\mathring{p}^s$ to $\mathring{p}'$, and update $\mathring{p}^s$ to $\mathring{p}'$.
- let $\mathring{p}'' = \left(\mathring{p}^e_{-k_i}, p^{(\beta_i,t]}_{k_i,\min}\right)$; in P^e, connect $\mathring{p}^e$ to $\mathring{p}''$, and update $\mathring{p}^e$ to $\mathring{p}''$.

(d) If $p^{(\beta_i,t]}_{k_i,\min} \le l_i \le p^{(\beta_i,t]}_{k_i,\max} < r_i$, let $\mathring{p}' = \left(\mathring{p}^e_{-k_i}, p^{(\beta_i,t]}_{k_i,\max}\right)$; in P^e, connect $\mathring{p}^e$ to $\mathring{p}'$, and update $\mathring{p}^e$ to $\mathring{p}'$.

(e) If $p^{(\beta_i,t]}_{k_i,\min} \le p^{(\beta_i,t]}_{k_i,\max} < l_i \le r_i$,
- let $\mathring{p}' = \left(\mathring{p}^s_{-k_i}, p^{(\beta_i,t]}_{k_i,\max}\right)$; in P^s, connect $\mathring{p}^s$ to $\mathring{p}'$, and update $\mathring{p}^s$ to $\mathring{p}'$.
- let $\mathring{p}'' = \left(\mathring{p}^e_{-k_i}, p^{(\beta_i,t]}_{k_i,\max}\right)$; in P^e, connect $\mathring{p}^e$ to $\mathring{p}''$, and update $\mathring{p}^e$ to $\mathring{p}''$.

3. Concatenate P^s and P^e at $\mathring{p}^s = \mathring{p}^e$ to form the path P.

There are at most $2m$ edges in the path P, with at most two edges added in each of the m steps. Note that the length of each edge added in the i-th step is at most $|\Delta p_{k_i}|$, so by simple calculus, the change to $\nabla_j(p')$ along each such edge is at most $H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right) \cdot |\Delta p_{k_i}|$. This yields (10).

To prove (4) and (5), first note that since $\tilde{P}^{[\beta_i,t]}_j\left(p^t_j\right) \subset \tilde{P}^{[\tau_j,t]}_j\left(p^t_j\right)$, $H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right) \le H^{[\tau_j,t]}_{k_ij}\left(p^t_j\right)$.

Then
$$\begin{aligned}\sum_{i=1}^m \frac{1}{\eta_i} H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right)\Delta t_{k_i} \le \sum_{i=1}^m \frac{1}{\eta_i} H^{[\tau_j,t]}_{k_ij}\left(p^t_j\right)\Delta t_{k_i} &\le \sum_{k\neq j} \frac{1}{\bar{\eta}_k} H^{[\tau_j,t]}_{kj}\left(p^t_j\right) \sum_{i:k_i=k} \Delta t_{k_i} \\ &\le 2\sum_{k\neq j} \frac{1}{\bar{\eta}_k} H^{[\tau_j,t]}_{kj}\left(p^t_j\right). \end{aligned} \tag{11}$$

The last inequality holds since $\sum_{i:k_i=k} \Delta t_{k_i} \le 1 + (t - \tau_j) \le 2$.

The proof of (4):

$$\begin{aligned}|\mu| \cdot (\tilde{g}_{j,\max} - \tilde{g}_{j,\min}) &\le 2\sum_{i=1}^m H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right) \cdot |\Delta p_{k_i}| \cdot |\mu| && \text{(by Eqn. (10))} \\ &\le \sum_{i=1}^m H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right) \cdot \left[\frac{\mu^2 \Delta t_{k_i}}{\eta_i} + \frac{\eta_i (\Delta p_{k_i})^2}{\Delta t_{k_i}}\right] && \text{(AM-GM ineq.)} \\ &\le 2\mu^2 \sum_{k\neq j} \frac{1}{\bar{\eta}_k} H^{[\tau_j,t]}_{kj}\left(p^t_j\right) + \sum_{i=1}^m \eta_i \cdot H^{[\beta_i,t]}_{k_ij}\left(p^t_j\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. && \text{(by Eqn. (11))}\end{aligned}$$

The proof of (5):

$$
\begin{aligned}
&\left(\tilde{g}_{j,\max} - \tilde{g}_{j,\min}\right)^2 \\
&\leq 4 \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \cdot \left|\Delta p_{k_{i_1}}\right| \cdot \left|\Delta p_{k_{i_2}}\right| \qquad \text{(by Eqn. (10))} \\
&\leq 2 \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \cdot \left[\frac{\left(\Delta p_{k_{i_1}}\right)^2 \eta_{i_1} \Delta t_{k_{i_2}}}{\eta_{i_2} \Delta t_{k_{i_1}}} + \frac{\left(\Delta p_{k_{i_2}}\right)^2 \eta_{i_2} \Delta t_{k_{i_1}}}{\eta_{i_1} \Delta t_{k_{i_2}}}\right] \qquad \text{(AM-GM ineq.)} \\
&= 2 \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \cdot \frac{\left(\Delta p_{k_{i_1}}\right)^2 \eta_{i_1} \Delta t_{k_{i_2}}}{\eta_{i_2} \Delta t_{k_{i_1}}} \\
&\qquad + 2 \sum_{i_2=1}^{m} \sum_{i_1=1}^{m} H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \cdot H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot \frac{\left(\Delta p_{k_{i_1}}\right)^2 \eta_{i_1} \Delta t_{k_{i_2}}}{\eta_{i_2} \Delta t_{k_{i_1}}} \\
&\qquad\qquad \text{(swap the indices } i_1 \text{ and } i_2 \text{ in the second double-summation)} \\
&= 4 \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \cdot \frac{\left(\Delta p_{k_{i_1}}\right)^2 \eta_{i_1} \Delta t_{k_{i_2}}}{\eta_{i_2} \Delta t_{k_{i_1}}} \\
&= 4 \left(\sum_{i_1=1}^{m} \eta_{i_1} \cdot H_{k_{i_1}j}^{[\beta_{i_1},t]}\left(p_j^t\right) \cdot \frac{\left(\Delta p_{k_{i_1}}\right)^2}{\Delta t_{k_{i_1}}}\right) \left(\sum_{i_2=1}^{m} \frac{1}{\eta_{i_2}} H_{k_{i_2}j}^{[\beta_{i_2},t]}\left(p_j^t\right) \Delta t_{k_{i_2}}\right) \\
&\leq 8 \left(\sum_{i=1}^{m} \eta_i \cdot H_{k_i j}^{[\beta_i,t]}\left(p_j^t\right) \cdot \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}\right) \left(\sum_{k \neq j} \frac{1}{\overline{\eta}_k} H_{kj}^{[\tau_j,t]}\left(p_j^t\right)\right). \qquad \text{(by Eqn. (11))}
\end{aligned}
$$

□

Proof of Lemma 5: First, we bound the integral terms in $\Phi(p^t, t, \tau)$ (see Eqn. (2)). Following the derivations of (7) and (8), with $\tilde{g}_j$ replaced by g_j, yields

$$
c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j}\, dt' \leq c_1(1 + 4\epsilon_{\mathrm{B}}) \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j} + c_1(2 + 8\epsilon_{\mathrm{B}}) \sum_{i=1}^{m} \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}
$$

and hence

$$
\sum_j c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j}\, dt' \leq c_1(1 + 4\epsilon_{\mathrm{B}}) \sum_j \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j} + c_1(2 + 8\epsilon_{\mathrm{B}}) \sum_j \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.
$$

When $2 - c_2 \geq c_1(2 + 8\epsilon_{\mathrm{B}})$, as $p_j^t = p_j^{\tau_j +}$, the double summation in the above inequality is no larger than the double summation in $\Phi(p^t, t, \tau)$. Thus $\Phi(p^t, t, \tau) \geq \phi(p^t) - c_1(1+4\epsilon_{\mathrm{B}}) \sum_j \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j}$.

Next, we bound the sum $\sum_j \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j}$. Suppose there are hypothetical updates to all the coordinates at time t, and p_j is updated with the most up-to-date gradient $\tilde{g}_j = g_j$ and step size $1/\gamma_j$. By Lemma 2 and Condition A2, $\phi^- - \phi^+ \geq \frac{1}{2} \sum_j \frac{(g_j)^2 \Delta t_j}{\gamma_j} \geq \frac{1}{2} \sum_j \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j}$. Here $\phi^- = \phi(p^t)$. Thus $\phi^- - \phi^+ \leq \phi(p^t) - \phi^* = \phi(p^t)$, and hence $\sum_j \frac{(g_j)^2 \Delta t_j}{\overline{\gamma}_j} \leq 2\phi(p^t)$. □

B Leontief Fisher Markets

Lemma 8. *Let τ_j, t be the times at which two consecutive updates to p_j occur. If γ_j^t is controlled and $c_2 \leq 1$, then $\Phi^{\tau_j+} - \Phi^{t+} \geq \left(1 - \frac{1}{\alpha} - 2\epsilon_B - 2\epsilon_F\right) \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j}$.*

Proof: This lemma can be proved by slightly modifying the proof of Lemma 4; we will use the notations defined therein.

By Lemma 4, Φ does not increase at the updates made in the time interval (τ_j, t). By (3),

$$\Phi^{\tau_j+} - \Phi^{t-} \geq c_1 \int_{\tau_j}^{t} \frac{(g_j(t'))^2}{\overline{\gamma}_j} \, dt' = E_2.$$

By (6),

$$\Phi^{t-} - \Phi^{t+} \geq \left(1 - \frac{1}{\alpha}\right) \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} - E_1 - E_2 + (2 - c_2) \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i, t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} - E_3.$$

Combining the two inequalities above yields

$$\begin{aligned} \Phi^{\tau_j+} - \Phi^{t+} &= \left(\Phi^{\tau_j+} - \Phi^{t-}\right) + \left(\Phi^{t-} - \Phi^{t+}\right) \\ &\geq \left(1 - \frac{1}{\alpha}\right) \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} + (2 - c_2) \sum_i \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i, t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} - E_1 - E_3. \end{aligned}$$

The result follows on noting that $p_j^{\tau_j+} = p_j^t$ and by applying the bounds on E_1 and E_3 in the proof of Lemma 4. $\square$

Let $U = \max\left\{\max_j\{p_j^0\}, 2\sum_i e_i\right\}$ be an upper bound on the prices throughout the tatonnement process [6].

Lemma 9. *Let $\delta = 1 - \frac{1}{\alpha} - 2\epsilon_B - 2\epsilon_F$. Suppose that there are consecutive updates to p_j at times $\Upsilon_0 < \Upsilon_1 < \cdots < \Upsilon_m$, where $\Upsilon_m - \Upsilon_0 \leq 2$. If $\left|p_j^{\Upsilon_0+} - p_j^{\Upsilon_m+}\right| \geq \epsilon$, where $\epsilon \leq 1$, then $\Phi^{\Upsilon_0+} - \Phi^{\Upsilon_m+} \geq \delta\epsilon^2 \cdot \min\left\{\frac{1}{2}, \frac{1}{3\lambda U}\right\}$.*

Proof: For $q = 1, 2, \cdots, m$, let $\Delta p_{j,q}$ be the change to p_j at the update timed Υ_q, and let $\tilde{z}_{j,q}$ be the $\tilde{z}$-value used for the update, i.e. $\gamma_j^{\Upsilon_q} = \frac{\max\{1, \tilde{z}_{j,q}\}}{\lambda p_j^{\Upsilon_q}}$ and $\Delta p_{j,q} = \lambda p_j^{\Upsilon_q} \cdot \min\{1, \tilde{z}_{j,q}\} \cdot \Delta t_q$.

We will use Lemma 8 to give a lower bound on the decrease to Φ between times Υ_0+ and Υ_m+. If $\tilde{z}_{j,q} < 1$, then

$$\frac{\gamma_j^{\Upsilon_q} (\Delta p_{j,q})^2}{\Delta t_q} = \frac{1}{\lambda p_j^{\Upsilon_q}} \frac{(\Delta p_{j,q})^2}{\Delta t_q} \geq \frac{1}{\lambda U} \frac{(\Delta p_{j,q})^2}{\Delta t_q}.$$

If $\tilde{z}_{j,q} \geq 1$, then

$$\frac{\gamma_j^{\Upsilon_q} (\Delta p_{j,q})^2}{\Delta t_q} = \frac{\tilde{z}_{j,q}}{\lambda p_j^{\Upsilon_q}} \cdot \lambda^2 \left(p_j^{\Upsilon_q}\right)^2 \Delta t_q = \lambda p_j^{\Upsilon_q} \tilde{z}_{j,q} \Delta t_q \geq |\Delta p_{j,q}|.$$

By Lemma 8,

$$\Phi^{\Upsilon_0+} - \Phi^{\Upsilon_m+} = \sum_{q=1}^{m} \left(\Phi^{\Upsilon_{q-1}+} - \Phi^{\Upsilon_q+}\right) \geq \delta \sum_{q=1}^{m} \frac{\gamma_j^{\Upsilon_q} (\Delta p_{j,q})^2}{\Delta t_q}$$
$$\geq \frac{\delta}{\lambda U} \sum_{q:\tilde{z}_{j,q}<1} \frac{(\Delta p_{j,q})^2}{\Delta t_q} + \delta \sum_{q:\tilde{z}_{j,q}\geq 1} |\Delta p_{j,q}|.$$

By the assumption $|p_j^{\Upsilon_0+} - p_j^{\Upsilon_m+}| \geq \epsilon$, $\sum_{q=1}^{m} |\Delta p_{j,q}| \geq \epsilon$. Let $\sigma := \epsilon^{-1} \sum_{q:\tilde{z}_{j,q}\geq 1} |\Delta p_{j,q}|$. Then $\sum_{q:\tilde{z}_{j,q}<1} |\Delta p_{j,q}| \geq \max\{0, (1-\sigma)\epsilon\}$. By the Cauchy-Schwarz inequality,

$$[\max\{0,(1-\sigma)\epsilon\}]^2 \leq \left(\sum_{q:\tilde{z}_{j,q}<1} |\Delta p_{j,q}|\right)^2 = \left(\sum_{q:\tilde{z}_{j,q}<1} \left|\frac{\Delta p_{j,q}}{\sqrt{\Delta t_q}}\right| \cdot \sqrt{\Delta t_q}\right)^2$$
$$\leq \left(\sum_{q:\tilde{z}_{j,q}<1} \frac{(\Delta p_{j,q})^2}{\Delta t_q}\right) \left(\sum_{q:\tilde{z}_{j,q}<1} \Delta t_q\right)$$
$$\leq 3 \sum_{q:\tilde{z}_{j,q}<1} \frac{(\Delta p_{j,q})^2}{\Delta t_q},$$

i.e. $\sum_{q:\tilde{z}_{j,q}<1} \frac{(\Delta p_{j,q})^2}{\Delta t_q} \geq \frac{1}{3} [\max\{0,(1-\sigma)\epsilon\}]^2$. Then

$$\Phi^{\Upsilon_0+} - \Phi^{\Upsilon_m+} \geq \frac{\delta}{3\lambda U} [\max\{0,(1-\sigma)\epsilon\}]^2 + \delta\sigma\epsilon.$$

The minimum value of the right hand side is at least $\delta\epsilon^2 \cdot \min\left\{\frac{1}{2}, \frac{1}{3\lambda U}\right\}$. □

Corollary 10. *For any $\epsilon > 0$, there exists a finite time T_ϵ such that for any good j, any $t \geq T_\epsilon$, and any $0 \leq \Delta t \leq 1$, $|p_j^t - p_j^{t+\Delta t}| \leq \epsilon$.*

Proof of Theorem 6 for the Leontief case: The proof comprises four steps. We need the following definitions: for any two price vectors p^A and p^B, let $d(p^A, p^B)$ denote the L_1 norm distance between the two price vectors, i.e. $d(p^A, p^B) = \sum_j |p_j^A - p_j^B|$. For any two sets of price vectors P^A and P^B, let $d(P^A, P^B) := \inf_{p^A \in P^A, p^B \in P^B} d(p^A, p^B)$.

Step 1. Let Ω be the set of limit points of a tatonnement process. We show that Ω is non-empty and connected.

Since all prices remain bounded by U throughout the tatonnement process, Ω is non-empty.

Suppose Ω is not connected. Let Ω_a denote a connected component of Ω, and let $\Omega_b = \Omega \backslash \Omega_a$. Suppose $d(\Omega_a, \Omega_b) = \epsilon' > 0$. By the definition of limit points, there exists a finite time $\Upsilon_{\epsilon'}$ such that thereafter the prices in the tatonnement process are always within an $\epsilon'/4$-neighborhood of either Ω_a or Ω_b. This forces an infinite number of updates, each separated by at least one time unit, such that each update makes a change to a price by at least at least $\epsilon'/2$. This contradicts Corollary 10.

Step 2. Recall that a market equilibrium is a price vector p^* at which for each j, $p^*_j > 0$ implies $z_j(p^*) = 0$ and $p^*_j = 0$ implies $z_j(p^*) \leq 0$. We define a pseudo-equilibrium: a price vector $\tilde{p}$ is a pseudo-equilibrium if for each j, $\tilde{p}_j > 0$ implies $z_j(\tilde{p}) = 0$. Note that every market equilibrium is a pseudo-equilibrium. We show that all limit points in Ω are pseudo-equilibria.

Suppose not. Let $p' \in \Omega$ be a price vector which is not a pseudo-equilibrium, i.e. there exists j such that $p'_j > 0$ but $z_j(p') \neq 0$. Let ϵ be a positive number such that for any price vector $\grave{p}$ in the ϵ-neighborhood of p', $\grave{p}_j \geq p'_j/2$ and $|z_j(\grave{p})| \geq |z_j(p')/2|$. By the definition of limit points, the tatonnement process enters the $(\epsilon/2)$-neighborhood of p' infinitely often. By Corollary 10, there exists a finite time such that subsequently, every time the tatonnement process enters the $\epsilon/2$-neighborhood of p', it stays in the ϵ-neighborhood of p' for at least one time unit. By Eqn. (3), Φ drops by at least $\lambda(p'_j/2)(z_j(p')/2)^2$ during each such stay in the ϵ-neighborhood of p'. This is a contradiction since Φ is positive throughout and hence cannot drop by at least $\lambda(p'_j/2)(z_j(p')/2)^2$ infinitely often.

Step 3. We show that the excess demands at all limit points in Ω are identical.

For every subset of goods S, let $\Omega_S = \{p' \in \Omega \,|\, p'_k > 0 \Leftrightarrow k \in S\}$. For each buyer, there are two cases:

- **if the buyer wants at least one good in S, say good ℓ:**
  Observe that by the definition of pseudo-equilibrium and Step 2, every price vector in Ω_S, excluding the zero prices in the price vector, is a market equilibrium for the sub-Leontief-market comprising the goods in S. Codenotti and Varadarajan [9] pointed out that the demands for the goods in S of each buyer are identical at every market equilibrium of the sub-Leontief market, and hence also in the original Leontief market. So the buyer demands the same positive but finite amount of good ℓ at every price vector in Ω_S in the original market. Also note that the buyer always demands the goods in the original market in a fixed proportion. This forces the demands for the goods not in S of the buyer are also identical at every price vector in Ω_S.

- **if the buyer wants no good in S:**
  Then the buyer demands infinite amount of each good that she wants, and demands zero amount of each good that she does not want.

In either case, the buyer's demands for each good at every price vector in Ω_S are identical, and hence also the total demand for each good.

Then consider a graph G with each vertex corresponding to a subset of goods S such that Ω_S is non-empty, and two vertices S_1, S_2 being adjacent if and only if $d\left(\Omega_{S_1}, \Omega_{S_2}\right) = 0$. Since excess demands are a continuous function[7] of prices, if S_1 and S_2 are adjacent, then the excess demands for all goods at every price vector in $S_1 \cup S_2$ are identical. By Step 1, the graph G is connected, thus the excess demands at all limit points in Ω are identical.

Step 4. We show that every limit point in Ω is indeed a market equilibrium.

[7]The range of the excess demand functions is the extended real line $\mathbb{R} \cup \{+\infty\}$; continuity of the excess demand function is w.r.t. the usual topology on the extended real line. To be specific, if $z_k(p) = +\infty$ for some p and k, then for any $M \in \mathbb{R}$, there exists an $\epsilon_M > 0$ such that $z_k(p) \geq M$ in the ϵ_M-neighborhood of p.

Suppose not, i.e. there exists a limit point p' in Ω which is a pseudo-equilibrium but not a market equilibrium, i.e. there exists k such that $p'_k = 0$ but $z_k(p') > 0$. By Step 3, z_k is positive at every limit point in Ω, and hence every p_k at every limit point must be zero. By the definition of limit points, for any $\epsilon > 0$, beyond a finite time, the tatonnement process must stay within the ϵ-neighborhood of Ω thereafter. By choosing a sufficiently small ϵ, z_k is bounded away from zero in the ϵ-neighborhood of Ω, and hence p_k increase indefinitely and eventually p_k becomes so large that the tatonnement process must leave the ϵ-neighborhood of Ω, a contradiction. $\square$

C Ongoing Complementary-CES Fisher Markets

The tatonnement process which we described in Section 5 is a two-stage process. In the first stage, the buyers repeatedly *report* their demands to sellers according to the current prices, then the sellers update the prices with the reported demands. The first stage continues until the market reaches a market equilibrium, and then trades occur in the second stage. Clearly, this is not a plausible real-world market dynamic.

In order to have a more realistic setting for a price adjustment algorithm, it would appear that out-of-equilibrium trade must be allowed, so as to generate the demand imbalances that then induce price adjustments. In an attempt to build a more realistic market model, Cole and Fleischer [10] introduced the *Ongoing market model*. In an ongoing Fisher market, the market repeats over an unbounded number of time intervals called *days*. Each day, the seller of each good receives one new unit of the good, and each buyer i is given e_i amount of money. In that day, each buyer i purchases a utility-maximizing bundle of goods of cost at most e_i.

But then there needs to be a way for seller to handle excess supply/demand. To this end, for each good j there is a *warehouse* of finite capacity χ_j which can meet excess demand and store excess supply. When there is surplus (supply exceeds demand), it is stored in the warehouse; when there is excess demand (demand exceeds supply), good is taken from the warehouse to meet the excess demand. The sellers change prices as needed to ensure their warehouses neither overfill nor run out of goods.

Given initial prices p^0, initial warehouses stocks v^0, where $0 < v_j^0 < \chi_j$ for each good j, and *ideal warehouse stocks* v^*, the task is to repeatedly adjust prices so as to converge to a market equilibrium with the warehouse stocks converging to their ideal values; for simplicity, we suppose that $v_j^* = \chi_j/2$ for each good j. v_j will denote the difference between the content of the warehouse of good j and v_j^*; hence $v_j \in [-\chi_j/2, \chi_j/2]$.

In an ongoing Fisher market, the sellers adjust the prices of their goods. In order to have progress, the sellers are required to update prices at least once per day. However, there is no upper bound on the frequency of price changes. This entails measuring demand on a finer scale than day units. Accordingly, we assume that each buyer spends their money at a uniform rate throughout the day, and hence *instantaneous demand* and *instantaneous excess demand* for good j at any time $t \in \mathbb{R}^+$ can be readily defined; we denote them by x_j^t and z_j^t respectively.

In this section, we analyse ongoing complementary-CES Fisher markets. Recall that for a complementary-CES Fisher market, tatonnement is equivalent to gradient descent on the convex function $\phi(p) = \sum_j p_j + \sum_i \hat{u}_i(p)$, where $\hat{u}_i(p)$ is the optimal utility that buyer i can attain at prices p. We will introduce new potential functions, which incorporate ϕ as a component, for the ongoing market analysis.

We use the following price update rule, which is a variant of (9), and which ensures convergence to the ideal warehouse stocks as well as to the market equilibrium:

$$p_j' = p_j \cdot \left(1 + \lambda_j \cdot \min\{\tilde{z}_j - \kappa_j v_j, 1\} \cdot \Delta t_j\right), \tag{12}$$

where λ_j, κ_j are small constants. Note that $\gamma_j^t = \frac{1}{p_j \lambda_j} \cdot \max\{1, \tilde{z}_j - \kappa_j v_j\}$.

Theorem 11. *If $\lambda_j \leq 1/60$ for all j, then there exists $\kappa_j > 0$ such that price updates using Rule* (12) *converge toward the market equilibrium in any complementary-CES Fisher market, with the warehouse stocks converging to their ideal values.*

First, we impose the following bounds on λ_j and κ_j.

B1. $\lambda_j \leq 1/60$;

B2. $\kappa_j/\lambda_j \leq 1/10$ (this, together with Condition B1, yields $\kappa_j \leq 1/600$);

B3. $|\kappa_j v_j| \leq 1/10$ always (such κ_j exist since the warehouse sizes are bounded).

We will impose more bounds on κ_j, but eventually we will show that, given any fixed λ_j satisfying Condition B1, for all j, there exist positive κ_j that satisfy all these bounds.

We need to be cautious with Condition B3, and also Condtion B4 which we will state later. At this point, it is not clear that v_j remains bounded throughout the tatonnement process, so the two conditions might cease to hold no matter how small κ_j is set. We show that this never happens in Section C.3.

Notations Let $f \geq 1$. A price vector p is f-bounded if, for all j, $\frac{1}{f} \leq \frac{p_j}{p_j^*} \leq f$. Let $R(f)$ denote the set of all f-bounded price vectors.

Our analysis comprises two phases. Phase 1 finishes when prices are guaranteed to be 1.9-bounded thereafter, and then we proceed to Phase 2. We outline the analysis of the two phases in Sections C.1 and C.2, respectively. We defer most proofs to Section C.4.

One component of the potential functions we will use is (similar to) Φ as defined in (2), and we will use some results from Sections 3 and 5. We deduce the values of $\epsilon_{\mathrm{B}}, \epsilon_{\mathrm{F}}$ that satisfy Conditions A3 and A4. Recall that by Property 3 of complementary-CES markets (see the appendix on tatonnement), if $\frac{\Delta p_j}{p_j} \leq 1/6$, then $\phi(p + \Delta p) - \phi(p) - \nabla_j \phi(p) \cdot \Delta p_j \leq \frac{1.5 x_j}{p_j}(\Delta p_j)^2$, where $x_j = z_j + 1$. Let $\tilde{x}_j = \tilde{z}_j + 1$. Recall that $x_j \leq \frac{100}{81}\tilde{x}_j \leq 1.24\tilde{x}_j$. Then

$$\frac{1.5x_j}{p_j} \cdot \frac{1}{\gamma_j^t} \leq \frac{1.86\tilde{x}_j}{p_j} \cdot \frac{p_j \lambda_j}{\max\{\tilde{z}_j - \kappa_j v_j, 1\}} \leq 1.86\lambda_j \cdot \frac{\tilde{z}_j + 1}{\max\{\tilde{z}_j - 0.1, 1\}} \leq \frac{1.86}{60} \cdot 2.1 < \frac{1}{15} < \frac{1}{2}, \quad (13)$$

and hence $\frac{\gamma_j^t}{15} > \frac{1.5x_j}{p_j}$. By Lemma 7 plus Conditions (A3) and (A4), we can set

$$\epsilon_{\mathrm{F}} = \frac{1.53}{1.5} \cdot \frac{1}{15} = 0.068 \quad (14)$$

$$\text{and} \quad \epsilon_{\mathrm{B}} = \frac{1.89}{1.5} \cdot \frac{1}{15} = 0.084. \quad (15)$$

Lemma 12. *Suppose there is an update to p_j at time t according to rule* (12). *Suppose that Conditions B1 and B3 hold. Let ϕ^- and ϕ^+ denote, respectively, the convex function values just before and just after the update. Let $z_j = -\nabla_j \phi(p^t)$ and $\tilde{z}_j \equiv \tilde{z}_j(t)$. Let Δp_j be the change to p_j made by the update, i.e. $\Delta p_j := \lambda_j p_j \cdot \min\{\tilde{z}_j - \kappa_j v_j, 1\} \cdot \Delta t_j$. Then*

$$\phi^- - \phi^+ \geq \frac{1}{2}\frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t} - \frac{1}{2}\frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \quad (16)$$

and

$$\phi^- - \phi^+ \geq \frac{41}{60}\frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} - \frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j|. \quad (17)$$

C.1 Phase 1

For Phase 1, we use the potential function $\Xi_1 \equiv \Xi_1(p^t, v^t, t, \tau)$:

$$\Xi_1 = \phi(p^t) - c_1 \sum_j \int_{\tau_j}^{t} \frac{(z_j(t'))^2}{\gamma_j^{\sigma_j}}\, dt' + \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} [2 - c_2(t-\beta_i)]$$
$$+ \sum_j \frac{(\kappa_j v_j^t)^2 (t-\tau_j)}{\gamma_j^{\sigma_j}}. \tag{18}$$

When there is no update, we show that

$$\frac{d\Xi_1}{dt} \le -\sum_j (c_1 - \kappa_j) \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} + \sum_j (1+\kappa_j) \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. \tag{19}$$

When there is an update, we show that

Lemma 13. *Suppose that there is an update to p_j at time t. Suppose that Conditions B1 and B3 hold. Let Ξ_1^- and Ξ_1^+, respectively, denote the values of Ξ_1 just before and just after the update. Then*

$$\Xi_1^- - \Xi_1^+ \ge \left(\frac{1}{4} - 1.4c_1\right) \frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t} + (1 - c_2 - 2.7c_1) \sum_{i=1}^{m} \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

Thus, by setting $c_1 = 5/28$ and $c_2 = 1/2$, Ξ_1 does not increase at any update.

Since ϕ is strongly convex, in the proof of Theorem 1(b), we show that $\sum_j \frac{(z_j)^2}{\overline{\gamma}_j} \ge D_1 \cdot \phi(p^t)$ for some positive constant $D_1 \le 1/10$. Let $\psi := \frac{1}{1.001} \cdot \inf_{p' \notin R(1.9)} \phi(p')$. We impose an additional condition on κ_j:

B4. κ_j are sufficiently small such that $\sum_j \frac{(\kappa_j v_j)^2}{\gamma_j^t} \le \frac{1}{26/D_1+4}\psi$ always.

Lemma 14. *If Condition B4 holds and $\Xi_1 \ge \psi/2$, then $\frac{d\Xi_1}{dt} \le -\Theta(1) \cdot \Xi_1(t)$.*

Proof: Let $H(t)$ denote the sum $\sum_j \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}$ at time t. By (18) and Condition B4,

$$\phi(p^t) + 2H(t) + \frac{1}{26/D_1+4}\psi \ \ge\ \Xi_1(t) \ \ge\ \psi/2.$$

Hence

$$\phi(p^t) + 2H(t) \ \ge\ \left(\frac{1}{2} - \frac{1}{26/D_1+4}\right)\psi \tag{20}$$

$$\text{and} \quad \phi(p^t) + 2H(t) \ \ge\ \left(1 - \frac{1}{13/D_1+2}\right)\Xi_1(t). \tag{21}$$

With our choices of c_1, c_2 and Condition B4, (19) yields

$$\begin{aligned}
\frac{d\Xi_1}{dt} &\le -\sum_j \left(\frac{5}{28} - \kappa_j\right) \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} + \sum_j (1+\kappa_j) \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} - \frac{1}{2} H(t) \\
&\le -\frac{1}{6} \sum_j \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} + \frac{601}{600} \cdot \frac{1}{26/D_1 + 4} \psi - \frac{1}{2} H(t) \\
&\le -\frac{D_1}{6} \cdot \phi(p^t) - \frac{1}{2} H(t) + \frac{601}{600} \cdot \frac{1}{26/D_1 + 4} \psi \\
&\le -\frac{D_1}{6} \big(\phi(p^t) + 2H(t)\big) + \frac{\frac{601}{600} \cdot \frac{1}{26/D_1+4}}{\frac{1}{2} - \frac{1}{26/D_1+4}} \big(\phi(p^t) + 2H(t)\big) && \text{(by Eqn. (20))} \\
&\le -\frac{D_1}{12} \big(\phi(p^t) + 2H(t)\big) \\
&\le -\frac{D_1}{12} \cdot \left(1 - \frac{1}{13/D_1 + 2}\right) \Xi_1(t) && \text{(by Eqn. (21))} \\
&\le -\frac{D_1}{13} \cdot \Xi_1(t). && (22)
\end{aligned}$$

□

Lemma 15. *If $\Xi_1(t_1) < \psi/2$ at some time t_1, then $\Xi_1(t) \le \psi/2$ thereafter.*

Proof: Suppose the contrary, i.e. at some time $t_2 > t_1$, $\Xi_1(t_2) > \psi/2$. Let T_2 be the collection of all such t_2, and let t' be the infimum of T_2. By Lemma 13 and our choices of c_1 and c_2, Ξ_1 never increases at an update. Hence, for Ξ_1 to exceed $\psi/2$ after time t_1, it must be due to continuous incrementing. This forces $\Xi_1(t') = \psi/2$ and $\frac{d\Xi_1}{dt}\big|_{t=t'} \ge 0$. But these contradict Lemma 14. □

Following the proof of Lemma 5, we obtain that $\Xi_1 \ge \phi(p^t) - 2c_1(1 + 8\epsilon_{\mathrm{B}})\phi(p^t)$, and as $c_1(1 + 8\epsilon_{\mathrm{B}}) \le \frac{1}{4}$, $\Xi_1 \ge \frac{1}{2}\phi(p^t)$. Thus if $\Xi_1 \le \psi/2$, then $\phi(p^t)/2 \le \Xi_1 \le \psi/2$. This implies $\phi(p^t) < \min_{p' \notin R(1.9)} \phi(p')$ and thus $p^t \in R(1.9)$. Lemma 14 shows that Ξ_1 decreases linearly until it drops below $\psi/2$ at some time t_1, and Lemma 15 shows that Ξ_1 remains below $\psi/2$ thereafter. Hence, $\forall t \ge t_1$, $p^t \in R(1.9)$ and we proceed to the analysis of Phase 2.

C.2 Phase 2

Phase 2 starts when all prices are guaranteed to be 1.9-bounded thereafter. Then each demand is between $\frac{1}{1.9}$ and 1.9 and hence $-0.5 \le z_j, \tilde{z}_j \le 0.9$. Since $|\kappa_j v_j| \le 0.1$ always, in Phase 2 the update rule (12) is equivalent to

$$p'_j = p_j \cdot (1 + \lambda_j \cdot (\tilde{z}_j - \kappa_j v_j) \cdot \Delta t_j), \tag{23}$$

i.e. $\gamma_j^t = \frac{1}{\lambda_j p_j}$.

In this phase, we will use a new potential function Ξ_2, which comprises two main components Φ and $\mathcal{W}$. Φ reflects how far the current prices are from the market equilibrium, and $\mathcal{W}$ accounts for the warehouse imbalances.

C.2.1 Component Φ

The first component of Ξ_2, $\Phi \equiv \Phi(p^t, t, \tau)$, is

$$\Phi = \phi(p^t) - c_1 \sum_j \int_{\tau_j}^t \lambda_j p_j (z_j(t'))^2 \, dt' + \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \left[6 - c_2(t - \beta_i)\right]. \tag{24}$$

When there is no update, it is straightforward to show that

$$\frac{d\Phi}{dt} = -c_1 \sum_j \lambda_j p_j (z_j^t)^2 - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. \tag{25}$$

When there is an update, we show that

Lemma 16. *Suppose that there is an update to p_j at time t. Suppose that Conditions B1 and B3 hold. Let Φ^- and Φ^+, respectively, denote the values of Φ just before and just after the update. Then*

$$\Phi^- - \Phi^+ \ge \left(\frac{1}{20} - 1.4 c_1\right) \lambda_j p_j (\tilde{z}_j)^2 \Delta t_j + 0.039 \frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j} - \frac{19}{20} \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j$$
$$+ (5 - c_2 - 2.7 c_1) \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

C.2.2 Component $\mathcal{W}$

Let $f_j := \ln(p_j / p_j^*)$. The second component of Ξ_2, $\mathcal{W} \equiv \mathcal{W}(p^t, v^t, t, \tau)$, is

$$\mathcal{W} = \sum_j \frac{\kappa_j}{\lambda_j} p_j^* \left(f_j + \lambda_j v_j\right)^2 - c_3 \sum_j \lambda_j p_j^* (\kappa_j v_j)^2 (t - \tau_j) + 2 \sum_j \kappa_j \lambda_j p_j^* \int_{\tau_j}^t v_j(t') z_j(t') \, dt'.$$

When there is no update, we show that for any $R_1 \in \mathbb{R}^+$,

$$\frac{d\mathcal{W}}{dt} \le -c_3 \sum_j (1 - \kappa_j) \lambda_j p_j^* (\kappa_j v_j^t)^2 + \sum_j (R_1 + c_3 \lambda_j) \kappa_j p_j^* (z_j^t)^2 + \frac{1}{R_1} \sum_j \kappa_j p_j^* (f_j)^2. \tag{26}$$

We will choose an appropriate value of R_1 at the end.

Lemma 17. *Suppose that there is an update to p_j at time t. Suppose that Conditions B1–B3 hold. Let $\mathcal{W}^-$ and $\mathcal{W}^+$, respectively, denote the values of $\mathcal{W}$ just before and just after the update. Then for any $R_2 \in \mathbb{R}^+$,*

$$\begin{aligned}\mathcal{W}^- - \mathcal{W}^+ \geq \left(0.858 - \frac{c_3}{1.9}\right) \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j - 0.0235 \lambda_j p_j (\tilde{z}_j)^2 \Delta t_j \\ - 3.809 \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \\ - 0.101 \kappa_j \frac{p_j^*(f_j)^2 \Delta t_j}{R_2} - 1.92 R_2 \frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j}.\end{aligned}$$

We will choose an appropriate value of R_2 at the end.

C.2.3 Ultimate Potential Function Ξ_2

The ultimate potential function $\Xi_2 \equiv \Xi_2(p^t, v^t, t, \tau)$ is

$$\Xi_2 := \Phi + 1.2\mathcal{W} + 0.1212 \sum_j \frac{\kappa_j p_j^*(f_j)^2}{R_2} (t - \tau_j).$$

From Lemmas 16 and 17, we deduce that

$$\begin{aligned}&(\Xi_2)^- - (\Xi_2)^+ \\ &\geq (0.039 - 2.304 R_2) \frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j} + (0.0218 - 1.4 c_1) \lambda_j p_j (\tilde{z}_j)^2 \Delta t_j + \left(0.0796 - \frac{12 c_3}{19}\right) \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j \\ &\quad + (0.4292 - c_2 - 2.7 c_1) \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. \qquad (27)\end{aligned}$$

From (25), (26) and the fact that $p_j \geq p_j^*/1.9$, we deduce that

$$\begin{aligned}\frac{d\Xi_2}{dt} \leq \sum_j \left[\frac{2.28 \kappa_j}{\lambda_j} (R_1 + c_3 \lambda_j) - c_1\right] \lambda_j p_j (z_j^t)^2 - 1.2 c_3 \sum_j (1 - \kappa_j) \lambda_j p_j^* (\kappa_j v_j^t)^2 \\ - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i, \sigma_j]} \left(p_j^{\tau_j +}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} + \left(\frac{1.2}{R_1} + \frac{0.1212}{R_2}\right) \sum_j \kappa_j p_j^* (f_j)^2. \qquad (28)\end{aligned}$$

We also show the following upper and lower bounds on Ξ_2.

If $\quad 2 - c_2 \geq 2.7 c_1, \qquad (29)$

$$\Xi_2 \geq (1 - 2.7 c_1) \phi(p^t) - 1.2 \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 - 20 \sum_j \kappa_j \lambda_j p_j (z_j)^2 + \sum_j \left(\frac{1}{5} - 1.2 c_3 \kappa_j\right) \kappa_j \lambda_j p_j^* (v_j)^2. \qquad (30)$$

Also,

$$\Xi_2 \le \phi(p^t) + \sum_j \left(\frac{2.4}{\lambda_j} + \frac{0.1212}{R_2}\right) \kappa_j p_j^*(f_j)^2 + 20 \sum_j \kappa_j \lambda_j p_j(z_j)^2$$
$$+ 10 \sum_j \sum_i \xi_j^{\beta_i} H_{kij}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} + 3.6 \sum_j \kappa_j \lambda_j p_j^*(v_j)^2. \tag{31}$$

In the next lemma, we show that $\sum_j p_j^*(f_j)^2 = O(1) \cdot \sum_j p_j(z_j)^2$, with the hidden constant in $O(1)$ depending on $\max_i \theta_i$, where θ_i is the parameter of the CES utility function of buyer i.

Lemma 18. *Let* $R := \left\{p' \mid \forall j,\ \frac{1}{1.9}p_j^* \le p_j' \le 1.9 p_j^*\right\}$ *and* $\bar\theta = \max_i \theta_i$.*For all* $p' \in R$,

$$\sum_j p_j^*(f_j)^2 \le \overline{M} \sum_j p_j'(z_j)^2,$$

where $\overline{M} = \left(1-\bar\theta\right)^{-1} \max\left\{26.56\ ,\ 6.64\bar\theta\left(1+\bar\theta-2^{\bar\theta}\right)^{-1}\right\}$.

Finally, we choose parameters R_1, R_2, c_1, c_2, c_3 such that Ξ_2 never increases at an update, and if there is no update, then $\frac{d\Xi_2}{dt} \le -\Theta(1)\cdot\Xi_2$. Set $R_2 = 39/2304$, $c_1 = \frac{0.0218}{1.4} \approx 0.0156$, $c_3 = \frac{19\times 0.0796}{12} \approx 0.1260$, $c_2 = 0.3855$ and $R_1 = 1$. By choosing sufficiently small κ_j, (28) and Lemma 18 yield

$$\frac{d\Xi_2}{dt} \le -\Theta(1)\cdot\sum_j \lambda_j p_j(z_j^t)^2 - \Theta\left(\min_j \kappa_j\right)\cdot\sum_j \kappa_j\lambda_j p_j^*(v_j^t)^2 - \Theta(1)\cdot\sum_j\sum_i \xi_j^{\beta_i} H_{kij}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

Also, by (31), Lemma 18 and the fact that $\phi(p) \le \Theta(1)\cdot\sum_j p_j(z_j)^2$ [6, Lemma 6.3] yield

$$\Xi_2 \le \Theta(1)\cdot\sum_j p_j(z_j^t)^2 + \Theta(1)\cdot\sum_j \kappa_j\lambda_j p_j^*(v_j^t)^2 + \Theta(1)\cdot\sum_j\sum_i \xi_j^{\beta_i} H_{kij}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

Thus $\frac{d\Xi_2}{dt} \le -\Omega\left(\min_j \kappa_j\right)\cdot\Xi_2$.

Further, (30) and the fact that $\phi(p) \ge \Theta(1)\cdot\sum_j p_j(z_j)^2$ [6, Lemma 6.2] yield

$$\Xi_2 \ge \Theta(1)\cdot\phi(p^t) + \Theta(1)\cdot\sum_j \kappa_j\lambda_j p_j^*(v_j)^2. \tag{32}$$

This implies that $\left(\phi(p^t) + \sum_j \kappa_j\lambda_j p_j^*(v_j)^2\right)$ decreases linearly, and finishes the proof of Theorem 11, except that we need to show Conditions B3 and B4 hold throughout the tatonnement process.

C.3 Warehouse Stocks Are Bounded

So far we need κ_j to satisfy Conditions B2, B3 and B4. Conditions B2 is satisfied so long as κ_j is sufficiently small. However, we need to be cautious with Conditions B3 and B4 as it is not immediately evident that v_j remains bounded throughout the tatonnement process.

We begin with Phase 1. The initial value of Ξ_1 decreases as κ_j decreases, and Phase 1 ends when Ξ_1 is smaller than $\psi/2$, which is independent of κ_j. By (22), Ξ_1 drops linearly at a rate that does not depend on κ_j. Hence, the length of Phase 1 is finitely bounded when the κ_j are sufficiently small. The change to each warehouse j is upper bounded by

$$\text{(The length of Phase 1)} \times \text{(Maximum excess demand for good } j \text{ in Phase 1)},$$

which is also finitely bounded. This allows us to set κ_j sufficiently small to ensure that Conditions B3 and B4 hold throughout Phase 1.

Next, we consider Phase 2, which starts at some time t_2. At t_2, which is the finishing time of Phase 1, Conditions B3 and B4 hold. Let $B := \Xi_2(t_2)$. Note that by (32), when Conditions B1–B4 hold, there exist constants C_1, C_2 such that

$$\Xi_2(t) \geq C_1 \phi(p^t) + C_2 \sum_j \kappa_j \lambda_j p_j^* (v_j)^2. \tag{33}$$

We impose two additional conditions on κ_j:

B5. κ_j are sufficiently small such that for all j, $\kappa_j \leq \frac{C_2 p_j^* \lambda_j}{101B}$.

B6. κ_j are sufficiently small such that for all j, $\kappa_j \leq \frac{C_2 \psi}{2(26/D_1+4)B}$.

Suppose that at some time $t_3 > t_2$, Condition B3 or B4 ceases to hold. By our analysis of Phase 2, Ξ_2 decreases between times t_2 and t_3, so $\Xi_2(t_3) \leq B$.

If Condition B3 ceases to hold at t_3, as the warehouse contents change smoothly, there exists a good ℓ with $|\kappa_\ell v_\ell| = 1/10$, and for other goods Condition B3 remains valid. Thus we can still apply (33) with Condition B5 to yield

$$\Xi_2(t_3) \geq C_2 \kappa_\ell \lambda_\ell p_\ell^* (v_\ell)^2 = \frac{C_2 \lambda_\ell p_\ell^*}{\kappa_\ell} |\kappa_\ell v_\ell|^2 = \frac{C_2 \lambda_\ell p_\ell^*}{100 \kappa_\ell} > B,$$

which is a contradiction.

If Condition B4 ceases to hold at t_3, as the warehouse contents change smoothly, $\sum_j p_j \lambda_j (\kappa_j v_j)^2 = \frac{1}{26/D_1+4} \psi$. Thus we can still apply (33) with Condition B6 to yield

$$\Xi_2(t_3) \geq C_2 \sum_j \kappa_j \lambda_j p_j^* (v_j)^2 \geq \frac{C_2}{1.9 \max_j \kappa_j} \sum_j p_j \lambda_j (\kappa_j v_j)^2 \geq \frac{C_2}{1.9 \kappa_j} \cdot \frac{1}{26/D_1 + 4} \psi > B,$$

which is a contradiction.

Thus, there does not exist $t_3 > t_2$ at which Condition B3 or B4 ceases to hold, i.e. the two conditions hold throughout Phase 1 and Phase 2.

C.4 Missing Proofs

Proof of Lemma 12: We start with the proof of (16). By Result (3) about Complementary CES markets (see the appendix on tatonnment):

$$
\begin{aligned}
\phi^- - \phi^+ \;&\geq\; [\tilde{z}_j + (z_j - \tilde{z}_j)](\Delta p_j) - \frac{1.5x_j}{p_j}(\Delta p_j)^2 \\
&\geq \tilde{z}_j(\Delta p_j) - \frac{1.5x_j}{p_j}(\Delta p_j)^2 - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \qquad (34)\\
&= \tilde{z}_j \frac{(\tilde{z}_j - \kappa_j v_j)\Delta t_j}{\gamma_j^t} - \frac{1.5x_j}{p_j}\left(\frac{(\tilde{z}_j - \kappa_j v_j)\Delta t_j}{\gamma_j^t}\right)^2 - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \\
&\geq \tilde{z}_j \frac{(\tilde{z}_j - \kappa_j v_j)\Delta t_j}{\gamma_j^t} - \frac{1}{2}\frac{(\tilde{z}_j - \kappa_j v_j)^2\Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \qquad \text{(By Eqn. (13) and } \Delta t_j \leq 1)\\
&= \frac{1}{2}\frac{(\tilde{z}_j)^2\Delta t_j}{\gamma_j^t} - \frac{1}{2}\frac{(\kappa_j v_j)^2\Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j|.
\end{aligned}
$$

Next, we give the proof of (17). From (34):

$$
\begin{aligned}
\phi^- - \phi^+ \;&\geq\; (\tilde{z}_j - \kappa_j v_j)(\Delta p_j) - \frac{1.5x_j}{p_j}(\Delta p_j)^2 - |z_j - \tilde{z}_j| \cdot |\Delta p_j| - |\kappa_j v_j| \cdot |\Delta p_j| \\
&\geq \frac{\gamma_j^t \Delta p_j}{\Delta t_j} \cdot \Delta p_j - \frac{1.5x_j}{p_j}\frac{1}{\gamma_j^t} \cdot \gamma_j^t(\Delta p_j)^2 - |z_j - \tilde{z}_j| \cdot |\Delta p_j| - \frac{1}{2}\left(2\frac{(\kappa_j v_j)^2\Delta t_j}{\gamma_j^t} + \frac{1}{2}\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j}\right) \\
&\qquad\qquad \text{(For the last term use the AM-GM ineq.)} \\
&\geq \frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - \frac{1}{15}\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - \frac{(\kappa_j v_j)^2\Delta t_j}{\gamma_j^t} - \frac{1}{4}\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \\
&\qquad\qquad \text{(For the second term use Eqn. (13) and } \Delta t_j \leq 1) \\
&= \frac{41}{60}\frac{\gamma_j^t(\Delta p_j)^2}{\Delta t_j} - \frac{(\kappa_j v_j)^2\Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j|.
\end{aligned}
$$

□

Proof of Equation (19)**:** Note that $\frac{dv_j}{dt} = -z_j^t$.

$$\begin{aligned}
&\frac{d\Xi_1}{dt} \\
&= -c_1 \sum_j \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} + \sum_j \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} - 2\sum_j \frac{(\kappa_j)^2 v_j^t z_j^t (t-\tau_j)}{\gamma_j^{\sigma_j}} \\
&\le -c_1 \sum_j \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} + \sum_j \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} + 2\sum_j \frac{\kappa_j}{\gamma_j^{\sigma_j}} \left|\kappa_j v_j^t\right| \cdot |z_j^t| \\
&\le -c_1 \sum_j \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} + \sum_j \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} + \sum_j \frac{\kappa_j}{\gamma_j^{\sigma_j}} \left[(\kappa_j v_j^t)^2 + (z_j^t)^2\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{(For the last term use the AM-GM ineq.)} \\
&= -\sum_j (c_1 - \kappa_j) \frac{(z_j^t)^2}{\gamma_j^{\sigma_j}} + \sum_j (1+\kappa_j) \frac{(\kappa_j v_j^t)^2}{\gamma_j^{\sigma_j}} - c_2 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.
\end{aligned}$$

□

Proof of Lemma 13:

$$\begin{aligned}
\Xi_1^- - \Xi_1^+ &= \phi^- - \phi^+ - c_1 \int_{\tau_j}^t \frac{(z_j(t'))^2}{\gamma_j^t}\, dt' + \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} [2 - c_2(t-\beta_i)] \\
&\qquad - 2\sum_{k\ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j} + \frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} \\
&\ge \frac{1}{2}\left(\frac{41}{60} \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} - \frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j|\right) \qquad \text{(By Eqn. (17))} \\
&\qquad + \frac{1}{2}\left(\frac{1}{2}\frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t} - \frac{1}{2}\frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} - |z_j - \tilde{z}_j| \cdot |\Delta p_j|\right) \qquad \text{(By Eqn. (16))} \\
&\qquad - c_1 \int_{\tau_j}^t \frac{(z_j(t'))^2}{\gamma_j^t}\, dt' + (2 - c_2) \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \\
&\qquad - 2\sum_{k\ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j} + \frac{(\kappa_j v_j)^2 \Delta t_j}{\gamma_j^t} \\
&\ge \frac{41}{120} \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} + \frac{1}{4}\frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t} - \underbrace{|z_j - \tilde{z}_j| \cdot |\Delta p_j|}_{F_1} - \underbrace{c_1 \int_{\tau_j}^t \frac{(z_j(t'))^2}{\gamma_j^t}\, dt'}_{F_2} \\
&\qquad + (2 - c_2) \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} - \underbrace{2\sum_{k\ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j}}_{F_3}.
\end{aligned}$$

Note that F_1, F_2 and F_3 are similar to the terms E_1, E_2 and E_3 in the proof of Lemma 4. We can bound F_1, F_2, F_3 similarly to the way we bounded E_1, E_2, E_3.

Recall from the proof of Lemma 4 that $V_2 := \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}$. We derive the following bounds:

$$F_1 \le 2\epsilon_{\mathrm{B}} \gamma_j^t (\Delta p_j)^2 + V_2;$$
$$F_2 \le c_1(1+4\epsilon_{\mathrm{B}}) \frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t} + c_1(2+8\epsilon_{\mathrm{B}}) V_2;$$
$$F_3 \le 2\epsilon_{\mathrm{F}} \gamma_j^t \frac{(\Delta p_j)^2}{\Delta t_j}.$$

Thus

$$\Xi_1^- - \Xi_1^+ \ge \left(\frac{41}{120} - 2\epsilon_{\mathrm{B}} - 2\epsilon_{\mathrm{F}}\right) \frac{\gamma_j^t (\Delta p_j)^2}{\Delta t_j} + \left(\frac{1}{4} - c_1(1+4\epsilon_{\mathrm{B}})\right) \frac{(\tilde{z}_j)^2 \Delta t_j}{\gamma_j^t}$$
$$+ \left(1 - c_2 - c_1(2+8\epsilon_{\mathrm{B}})\right) V_2.$$

Note that by Eqns. (14) and (15), $2\epsilon_{\mathrm{F}} + 2\epsilon_{\mathrm{B}} = 0.304 < \frac{41}{120}$, $1 + 4\epsilon_{\mathrm{B}} < 1.4$ and $2 + 8\epsilon_{\mathrm{B}} < 2.7$. The result now follows. □

Proof of Lemma 16: This proof is similar to the one of Lemma 13; we only point out the key steps.

$$\begin{aligned}
\Phi^- - \Phi^+ &\ge \phi^- - \phi^+ - c_1 \int_{\tau_j}^t \lambda_j p_j (z_j(t'))^2 \, dt' + (6 - c_2) \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \\
&\quad - 6 \sum_{k \ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j} \\
&\ge \frac{9}{10} \left(\frac{41}{60} \frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j} - \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \right) \qquad \text{(By Eqn. (17))} \\
&\quad + \frac{1}{10} \left(\frac{1}{2} \lambda_j p_j (\tilde{z}_j)^2 \Delta t_j - \frac{1}{2} \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j - |z_j - \tilde{z}_j| \cdot |\Delta p_j| \right) \qquad \text{(By Eqn. (16))} \\
&\quad - c_1 \int_{\tau_j}^t \lambda_j p_j (z_j(t'))^2 \, dt' + (6 - c_2) \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \\
&\quad - 6 \sum_{k \ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j} \\
&\ge \frac{123}{200} \frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j} + \frac{1}{20} \lambda_j p_j (\tilde{z}_j)^2 \Delta t_j - \frac{19}{20} \lambda_j p_j (\kappa_j v_j)^2 \Delta t_j - \underbrace{|z_j - \tilde{z}_j| \cdot |\Delta p_j|}_{F_1} \\
&\quad \underbrace{- c_1 \int_{\tau_j}^t \lambda_j p_j (z_j(t'))^2 \, dt'}_{F_2} + (6 - c_2) \sum_{i=1}^m \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,t]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \\
&\quad \underbrace{- 6 \sum_{k \ne j} \xi_k^t \cdot H_{jk}^{[t,\sigma_k]} \left(p_k^{\tau_k+}\right) \frac{(\Delta p_j)^2}{\Delta t_j}}_{F_3'}.
\end{aligned}$$

Then we apply the bounds on F_1, F_2, F_3 in the proof of Lemma 13.[8] to show that

$$\Phi^- - \Phi^+ \geq \left(\frac{123}{200} - 2\epsilon_{\mathrm{B}} - 6\epsilon_{\mathrm{F}}\right)\frac{(\Delta p_j)^2}{\lambda_j p_j \Delta t_j} + \left(\frac{1}{20} - c_1(1+4\epsilon_{\mathrm{B}})\right)\lambda_j p_j (\tilde{z}_j)^2 \Delta t_j$$
$$- \frac{19}{20}\lambda_j p_j (\kappa_j v_j)^2 \Delta t_j + (5 - c_2 - c_1(2+8\epsilon_{\mathrm{B}}))\, V_2.$$

Note that $\frac{123}{200} - 2\epsilon_{\mathrm{B}} - 6\epsilon_{\mathrm{F}} = 0.039$, $1+4\epsilon_{\mathrm{B}} < 1.4$ and $2+8\epsilon_{\mathrm{B}} < 2.7$; the lemma now follows. $\square$

Proof of Equation (26)**:** Note that $\frac{dv_j}{dt} = -z_j^t$.

$$\begin{aligned}
\frac{d\mathcal{W}}{dt} &= \sum_j p_j^* \left[\frac{2\kappa_j}{\lambda_j}(f_j + \lambda_j v_j^t)(-\lambda_j z_j^t) - c_3\lambda_j(\kappa_j v_j^t)^2 + 2c_3\lambda_j(\kappa_j)^2 v_j^t z_j^t (t-\tau_j) + 2\kappa_j\lambda_j v_j^t z_j^t\right] \\
&\leq \sum_j p_j^* \left[2\kappa_j |f_j||z_j^t| - c_3\lambda_j(\kappa_j v_j^t)^2 + 2c_3\lambda_j\kappa_j|\kappa_j v_j^t||z_j^t|\right] \\
&\leq \sum_j p_j^* \left[\kappa_j\left(\frac{(f_j)^2}{R_1} + R_1(z_j^t)^2\right) - c_3\lambda_j(\kappa_j v_j^t)^2 + c_3\lambda_j\kappa_j\left[(\kappa_j v_j^t)^2 + (z_j^t)^2\right]\right] \\
&= -c_3\sum_j(1-\kappa_j)\lambda_j p_j^*(\kappa_j v_j^t)^2 + \sum_j (R_1 + c_3\lambda_j)\kappa_j p_j^*(z_j^t)^2 + \frac{1}{R_1}\sum_j \kappa_j p_j^*(f_j)^2.
\end{aligned}$$

$\square$

Proof of Lemma 17: At the price update, $f_j^+ = f_j^- + \ln(1 + \lambda_j(\tilde{z}_j - \kappa_j v_j)\Delta t_j)$. Note that in Phase 2, $|\lambda_j(\tilde{z}_j - \kappa_j v_j)\Delta t_j| \leq 1/60$ and hence $\ln(1 + \lambda_j(\tilde{z}_j - \kappa_j v_j)\Delta t_j) = (1+\chi)\lambda_j(\tilde{z}_j - \kappa_j v_j)\Delta t_j$ for some χ with $|\chi| \leq \frac{1}{100}$.[9] Then

$$\mathcal{W}^- - \mathcal{W}^+ = p_j^*\left[\frac{\kappa_j}{\lambda_j}\left[(f_j + \lambda_j v_j)^2 - (f_j + (1+\chi)\lambda_j(\tilde{z}_j - \kappa_j v_j)\Delta t_j + \lambda_j v_j)^2\right]\right.$$
$$\left. - c_3\lambda_j(\kappa_j v_j)^2\Delta t_j + 2\kappa_j\lambda_j\int_{\tau_j}^t v_j(t')z_j(t')\,dt'\right]$$

Let $\bar{z}_j$ be the average excess demand for good j between times τ_j and t, i.e. $\bar{z}_j := \frac{1}{t_2-t_1}\int_{t_1}^{t_2} z_j^{t'}\,dt'$. Note that $v_j(\tau_j) = v_j(t) + \bar{z}_j\Delta t_j$ and $\frac{dv_j}{dt} = -z_j$. We use integration by substitution to evaluate the integral in the above formula:

$$\int_{\tau_j}^t v_j(t')z_j(t')\,dt' = -\int_{v_j(\tau_j)}^{v_j(t)} v_j\,dv_j = \frac{1}{2}\left(v_j(\tau_j)^2 - v_j(t)^2\right) = v_j\bar{z}_j\Delta t_j + \frac{1}{2}(\bar{z}_j)^2(\Delta t_j)^2.$$

[8]There is one minor difference: $\gamma_j^{\sigma_j}$ is replaced by $1/(\lambda_j p_j)$. Also, F_3' is three times the value of F_3, so the bound on F_3' is amplified accordingly.

[9]When $|y| \leq \frac{1}{60}$, $\ln(1+y) \in \left[1 - \frac{1}{100}, 1 + \frac{1}{100}\right]\cdot y$.

By direct expansion and regrouping terms, we have

$$
\begin{aligned}
&\mathcal{W}^- - \mathcal{W}^+ \\
&= p_j^* \Delta t_j \left\{ \left[2(1+\chi) - (1+\chi)^2 \kappa_j \Delta t_j - c_3\right] \lambda_j (\kappa_j v_j)^2 + \left[(\bar{z}_j)^2 - (\tilde{z}_j)^2\right] \kappa_j \lambda_j \Delta t_j \right. \\
&\qquad - (2\chi + \chi^2)\kappa_j \lambda_j (\tilde{z}_j)^2 \Delta t_j + 2\lambda_j (\bar{z}_j - \tilde{z}_j)\kappa_j v_j \\
&\qquad \left. + \left[(1+\chi)^2 \kappa_j \Delta t_j - \chi\right] \cdot 2\lambda_j \tilde{z}_j \kappa_j v_j - 2(1+\chi)\kappa_j f_j (\tilde{z}_j - \kappa_j v_j) \right\} \\
&\geq p_j^* \Delta t_j \left\{ \left[2(1+\chi) - (1+\chi)^2 \kappa_j \Delta t_j - c_3\right] \lambda_j (\kappa_j v_j)^2 - \underbrace{\kappa_j \lambda_j |(\bar{z}_j)^2 - (\tilde{z}_j)^2|}_{G_1} \right. \\
&\qquad - \underbrace{|2\chi + \chi^2| \cdot \kappa_j \lambda_j (\tilde{z}_j)^2}_{G_2} - \underbrace{2|\bar{z}_j - \tilde{z}_j| \cdot |\lambda_j \kappa_j v_j|}_{G_3} \\
&\qquad \left. - \underbrace{2|(1+\chi)^2 \kappa_j \Delta t_j - \chi| \lambda_j |\tilde{z}_j| \cdot |\kappa_j v_j|}_{G_4} - \underbrace{2(1+\chi)\kappa_j |f_j| \cdot |\tilde{z}_j - \kappa_j v_j|}_{G_5} \right\}
\end{aligned}
$$

Next, we bound the terms G_1, G_2, G_3, G_4, G_5. Recall the notations we use in the proof of Lemma 4 $V_1 := \sum_{k \neq j} \frac{1}{\min_{i:k_i=k} \xi_j^{\beta_i}} H_{kj}^{[\tau_j,t]}\left(p_j^t\right)$ and $V_2 := \sum_{i=1}^m \xi_j^{\beta_i} \cdot H_{k_i j}^{[\beta_i,t]}\left(p_j^t\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}$.

$$
\begin{aligned}
G_1 &\leq \kappa_j \lambda_j \left[(\bar{z}_j - \tilde{z}_j)^2 + \frac{2}{\lambda_j p_j} |\lambda_j p_j \tilde{z}_j| \cdot |\bar{z}_j - \tilde{z}_j|\right] \\
&\leq \kappa_j \lambda_j \left[8 V_1 V_2 + \frac{2}{\lambda_j p_j}\left(2(\lambda_j p_j \tilde{z}_j)^2 V_1 + V_2\right)\right] && \text{(By Eqns. (5) and (4))} \\
&\leq \kappa_j \lambda_j \left[8 \frac{\epsilon_{\mathrm{B}}}{\lambda_j p_j} V_2 + 4\epsilon_{\mathrm{B}} (\tilde{z}_j)^2 + \frac{2}{\lambda_j p_j} V_2\right] && \text{(as by Cond. (A2), } V_1 \leq \epsilon_{\mathrm{B}} \gamma_j^t = \epsilon_{\mathrm{B}}/(\lambda_j p_j)) \\
&= 4\epsilon_{\mathrm{B}} \kappa_j \lambda_j (\tilde{z}_j)^2 + \frac{(2 + 8\epsilon_{\mathrm{B}})\kappa_j}{p_j} V_2.
\end{aligned}
$$

To bound G_2, note that $|\chi| \leq 1/100$ and $\kappa_j \leq 1/600$ imply that $\kappa_j |2\chi + \chi^2| \leq 0.0000335$, and hence $G_2 \leq 0.0000335 \lambda_j (\tilde{z}_j)^2$.

$$
\begin{aligned}
G_3 &= \frac{2}{p_j} |\bar{z}_j - \tilde{z}_j| \cdot |\lambda_j p_j \kappa_j v_j| \\
&\leq \frac{2}{p_j}\left[2(\lambda_j p_j \kappa_j v_j)^2 V_1 + V_2\right] && \text{(By Eqn. (4))} \\
&\leq \frac{2}{p_j}\left[2(\lambda_j p_j \kappa_j v_j)^2 \frac{\epsilon_{\mathrm{B}}}{\lambda_j p_j} + V_2\right] \\
&= 4\epsilon_{\mathrm{B}} \lambda_j (\kappa_j v_j)^2 + \frac{2}{p_j} V_2.
\end{aligned}
$$

To bound G_4, note that $|\chi| \leq 1/100$, $\kappa_j \leq 1/600$ and $\Delta t_j \leq 1$ imply that $|(1+\chi)^2 \kappa_j \Delta t_j - \chi| \leq 0.0117$. Then by AM-GM inequality, $G_4 \leq 0.0117 \lambda_j (\tilde{z}_j)^2 + 0.0117 \lambda_j (\kappa_j v_j)^2$.

$$
\begin{aligned}
G_5 &= 2(1+\chi)\frac{\kappa_j}{\lambda_j p_j}|f_j| \cdot |\lambda_j p_j(\tilde{z}_j - \kappa_j v_j)| \\
&= 2(1+\chi)\frac{\kappa_j}{\lambda_j p_j}|f_j| \cdot \left|\frac{\Delta p_j}{\Delta t_j}\right| \\
&\le \frac{101}{100}\frac{\kappa_j}{\lambda_j p_j}\left(\frac{\kappa_j p_j (f_j)^2}{R_2} + \frac{R_2(\Delta p_j)^2}{\kappa_j p_j (\Delta t_j)^2}\right). \qquad \text{(by the AM-GM ineq.)}
\end{aligned}
$$

Combining all the above bounds yields

$$
\begin{aligned}
\mathcal{W}^- - \mathcal{W}^+ \ge{} & \left[2(1+\chi) - (1+\chi)^2\kappa_j\Delta t_j - c_3 - 4\epsilon_{\mathrm{B}} - 0.0117\right]\frac{p_j^*}{p_j}\lambda_j p_j(\kappa_j v_j)^2\Delta t_j \\
& - (0.0118 + 4\epsilon_{\mathrm{B}}\kappa_j)\frac{p_j^*}{p_j}\lambda_j p_j(\tilde{z}_j)^2\Delta t_j - \frac{(2+2\kappa_j+8\epsilon_{\mathrm{B}}\kappa_j)p_j^*}{p_j}V_2 \\
& - \frac{101}{100}\cdot\frac{\kappa_j}{\lambda_j}\cdot\frac{p_j^*}{p_j}\left(\frac{\kappa_j p_j (f_j)^2\Delta t_j}{R_2} + \frac{R_2(\Delta p_j)^2}{\kappa_j p_j \Delta t_j}\right).
\end{aligned}
$$

Note the following:

- $|\chi| \le 1/100$, $\kappa_j \le 1/600$ and $\Delta t_j \le 1$ imply that $2(1+\chi) - (1+\chi)^2\kappa_j\Delta t_j \ge 1.9783$. Also, recall that $\epsilon_{\mathrm{B}} = 0.084$. Thus $[2(1+\chi) - (1+\chi)^2\kappa_j\Delta t_j - c_3 - 4\epsilon_{\mathrm{B}} - 0.0117]\frac{p_j^*}{p_j} \ge (1.6306 - c_3)/1.9 \ge 0.858 - c_3/1.9$.
- $\epsilon_{\mathrm{B}} = 0.084$ and $\kappa_j \le 1/600$ imply that $(0.0118 + 4\epsilon_{\mathrm{B}}\kappa_j)\frac{p_j^*}{p_j} \le 0.01236 \times 1.9 \le 0.0235$.
- $\epsilon_{\mathrm{B}} = 0.084$ and $\kappa_j \le 1/600$ imply that $\frac{(2+2\kappa_j+8\epsilon_{\mathrm{B}}\kappa_j)p_j^*}{p_j} \le 2.00446 \times 1.9 \le 3.809$.
- $\frac{101}{100}\cdot\frac{p_j^*}{p_j} \le 1.92$.

The lemma follows. □

In the proofs of Equations (30) and (31) below, we need the following bound on $(\bar{z}_j)^2$:

$$
\begin{aligned}
(\bar{z}_j)^2 - (z_j)^2 &= (\bar{z}_j - z_j)^2 - 2z_j(z_j - \bar{z}_j) \\
&\le 8V_1V_2 + \frac{1}{5\lambda_j p_j}|10\lambda_j p_j z_j| \cdot |z_j - \bar{z}_j| \qquad \text{(by Eqn. 5)} \\
&\le \frac{8\epsilon_{\mathrm{B}}}{\lambda_j p_j}V_2 + \frac{1}{5\lambda_j p_j}\left(200(\lambda_j p_j)^2(z_j)^2V_1 + V_2\right) \qquad (\text{as } V_1 \le \epsilon_{\mathrm{B}}/(\lambda_j p_j)) \\
&\le \frac{0.672}{\lambda_j p_j}V_2 + 40\lambda_j p_j(z_j)^2\frac{\epsilon_{\mathrm{B}}}{\lambda_j p_j} + \frac{0.2}{\lambda_j p_j}V_2 \\
&= 3.36(z_j)^2 + \frac{0.872}{\lambda_j p_j}\sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}
\end{aligned}
$$

and hence

$$
\lambda_j p_j(\bar{z}_j)^2 \le 4.36\lambda_j p_j(z_j)^2 + 0.872\sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]}\left(p_j^{\tau_j+}\right)\frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}. \tag{35}
$$

Proof of Equation (30)**:** By Lemma 5, if $2 - c_2 \geq 2.7c_1$, then

$$\phi(p^t) - c_1 \sum_j \int_{\tau_j}^t \lambda_j p_j(z_j(t'))^2\, dt' + \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} [2 - c_2(t-\beta_i)] \geq (1-2.7c_1)\phi(p^t).$$

Thus, Φ, as defined in 24, satisfy

$$\begin{aligned}\Phi &= \phi(p^t) - c_1 \sum_j \int_{\tau_j}^t \lambda_j p_j(z_j(t'))^2\, dt' + \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} [6 - c_2(t-\beta_i)] \\ &\geq (1-2.7c_1)\phi(p^t) + 4 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.\end{aligned}$$

$$\begin{aligned}&\mathcal{W} \\ &= \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j + \lambda_j v_j)^2 - c_3 \sum_j \lambda_j p_j^* (\kappa_j v_j)^2 (t-\tau_j) + 2 \sum_j \kappa_j \lambda_j p_j^* \int_{\tau_j}^t v_j(t') z_j(t')\, dt' \\ &\geq \sum_j \frac{\kappa_j}{\lambda_j} p_j^* \left(\frac{(\lambda_j v_j)^2}{2} - (f_j)^2\right) - c_3 \sum_j \lambda_j p_j^* (\kappa_j v_j)^2 + 2 \sum_j \kappa_j \lambda_j p_j^* \left(v_j \bar{z}_j (t-\tau_j) + \frac{1}{2}(\bar{z}_j)^2 (t-\tau_j)^2\right) \\ &\geq \sum_j \left(\frac{1}{2} - c_3\kappa_j\right) \lambda_j \kappa_j p_j^* (v_j)^2 - \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 \\ &\qquad + 2 \sum_j \kappa_j \lambda_j p_j^* \left(-\frac{1}{6}(v_j)^2 - \frac{3}{2}(\bar{z}_j)^2 (t-\tau_j)^2 + \frac{1}{2}(\bar{z}_j)^2 (t-\tau_j)^2\right) \qquad \text{(by the AM-GM ineq.)} \\ &\geq \sum_j \left(\frac{1}{6} - c_3\kappa_j\right) \lambda_j \kappa_j p_j^* (v_j)^2 - \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 - 2 \sum_j \kappa_j \lambda_j p_j^* (\bar{z}_j)^2 \\ &\geq \sum_j \left(\frac{1}{6} - c_3\kappa_j\right) \lambda_j \kappa_j p_j^* (v_j)^2 - \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 \\ &\qquad - 3.8 \sum_j \kappa_j \left(4.36 \lambda_j p_j (z_j)^2 + 0.872 \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}\right) \qquad \text{(by eqn. (35))} \\ &\geq \sum_j \left(\frac{1}{6} - c_3\kappa_j\right) \lambda_j \kappa_j p_j^* (v_j)^2 - \sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 \\ &\qquad - 16.6 \sum_j \kappa_j \lambda_j p_j (z_j)^2 - 3.314 \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.\end{aligned}$$

Recall that $\Xi_2 = \Phi + 1.2\mathcal{W} + 0.1212 \sum_j \frac{\kappa_j p_j^* (f_j)^2}{R_2} (t-\tau_j) \geq \Phi + 1.2\mathcal{W}$. With the two inequalities above, the result follows. □

Proof of Equation (31)**:** It follows immediately from (24) that

$$\Phi \leq \phi(p^t) + 6 \sum_j \sum_i \xi_j^{\beta_i} H_{k_i j}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j+}\right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.$$

$$
\begin{aligned}
\mathcal{W} &= \sum_j \frac{\kappa_j}{\lambda_j} p_j^* \left(f_j + \lambda_j v_j\right)^2 - c_3 \sum_j \lambda_j p_j^* (\kappa_j v_j)^2 (t-\tau_j) + 2\sum_j \kappa_j \lambda_j p_j^* \int_{\tau_j}^{t} v_j(t') z_j(t')\, dt' \\
&\le 2\sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 + 2\sum_j \kappa_j\lambda_j p_j^* (v_j)^2 + 2\sum_j \kappa_j \lambda_j p_j^* \left(v_j \bar{z}_j (t-\tau_j) + \frac{1}{2}(\bar{z}_j)^2 (t-\tau_j)^2 \right) \\
&\le 2\sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 + 2\sum_j \kappa_j\lambda_j p_j^* (v_j)^2 + 2\sum_j \kappa_j \lambda_j p_j^* \left(\frac{1}{2}(v_j)^2 + \frac{1}{2}(\bar{z}_j)^2 (t-\tau_j)^2 + \frac{1}{2}(\bar{z}_j)^2 (t-\tau_j)^2 \right) \\
&\le 2\sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 + 3\sum_j \kappa_j\lambda_j p_j^* (v_j)^2 + 2\sum_j \kappa_j\lambda_j p_j^* (\bar{z}_j)^2 \\
&\le 2\sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 + 3\sum_j \kappa_j\lambda_j p_j^* (v_j)^2 \\
&\qquad + 3.8 \sum_j \kappa_j \left(4.36 \lambda_j p_j (z_j)^2 + 0.872 \sum_i \xi_j^{\beta_i} H_{kij}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j +} \right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}} \right) \qquad \text{(by eqn. (35))} \\
&\le 2\sum_j \frac{\kappa_j}{\lambda_j} p_j^* (f_j)^2 + 3\sum_j \kappa_j\lambda_j p_j^* (v_j)^2 \\
&\qquad + 16.6 \sum_j \kappa_j \lambda_j p_j (z_j)^2 + 3.314 \sum_i \xi_j^{\beta_i} H_{kij}^{[\beta_i,\sigma_j]} \left(p_j^{\tau_j +} \right) \frac{(\Delta p_{k_i})^2}{\Delta t_{k_i}}.
\end{aligned}
$$

Recall that $\Xi_2 = \Phi + 1.2\mathcal{W} + 0.1212 \sum_j \frac{\kappa_j p_j^* (f_j)^2}{R_2} (t-\tau_j) \le \Phi + 1.2\mathcal{W} + 0.1212 \sum_j \frac{\kappa_j p_j^* (f_j)^2}{R_2}$. With the two inequalities above, the result follows. $\square$

To prove Lemma 18, we need the following lemma.

Lemma 19. *For all $p' \in R(1.9)$, $\phi(p') \ge \frac{1-\bar{\theta}}{13.28} \sum_j p_j^* (f_j)^2$.*

Proof: Let $x_{ij}(p')$ be the demand for good j of buyer i at price p'. Note that

$$
\frac{\partial^2 \phi}{\partial (p_j)^2}(p') = \sum_i \left(\frac{\theta_i (x_{ij}(p'))^2}{e_i} + \frac{(1-\theta_i) x_{ij}(p')}{p'_j} \right) \quad \text{and} \quad \frac{\partial^2 \phi}{\partial p_j \partial p_k}(p') = \sum_i \frac{\theta_i x_{ij}(p') x_{ik}(p')}{e_i}.
$$

Let $A^i(p')$ denote the matrix with $A^i_{jk}(p') = x_{ij}(p') x_{ik}(p')$. Let $B^i(p')$ denote the diagonal matrix with $B^i_{jj}(p') = x_{ij}(p')/p'_j$. Then the Hessian of ϕ at p', which we denote it by $H(p')$, is $\sum_i \frac{\theta_i}{e_i} A^i(p') + \sum_i (1-\theta_i) B^i(p')$.

There are two key observations: first that A^i is positive semi-definite and second that $\sum_i (1-\theta_i) B^i(p')$ majorizes $(1-\bar{\theta}) \sum_i B^i(p')$, where $\bar{\theta} = \max_i \theta_i$. Hence $H(p')$ majorizes $(1-\bar{\theta}) \sum_i B^i(p') := (1-\bar{\theta}) B(p')$, where $B_{jj}(p') = x_j(p')/p'_j$. As $p' \in R(1.9)$, $x_j(p') \ge 1/1.9$ and $p'_j \le 1.9 p_j^*$. Hence $B_{jj}(p') \ge \frac{1}{3.61 p_j^*}$.

Next, consider the function $\bar{\phi}(p) = \phi(p) - \sum_j \frac{1-\bar{\theta}}{7.22 p_j^*} (p_j - p_j^*)^2$. Observe that for all j, $\frac{\partial \bar{\phi}}{\partial p_j}(p^*) = 0$ and the Hessian of $\bar{\phi}$ at every $p' \in R$ majorizes the zero matrix; consequently, $\bar{\phi}$

is convex in $R(1.9)$, and p^* is its minimum point. Note that $\bar{\phi}(p^*) = 0$, so for all $p' \in R(1.9)$, $\phi(p') \geq \sum_j \frac{1-\bar{\theta}}{7.22 p_j^*}(p_j' - p_j^*)^2$.

Since $\left(\frac{p_j' - p_j^*}{p_j^*}\right)^2 \geq 0.544 \ln^2 \frac{p_j'}{p_j^*} = 0.544(f_j)^2$, $\phi(p') - \phi^* \geq \frac{1-\bar{\theta}}{13.28} \sum_j p_j^*(f_j)^2$. □

Proof of Lemma 18: [6, Lemma 6.3] showed that for all $p' \in R(1.9)$, $\phi(p') \leq \max\left\{2, \frac{\bar{\theta}}{2\left(1+\bar{\theta}-2^{\bar{\theta}}\right)}\right\} \cdot \sum_j p_j'(z_j)^2$. Combining this with Lemma 19 yields the result. □