Data assimilation in price formation

Martin Burger* Jan-Frederik Pietschmann† Marie-Therese Wolfram‡

June 3, 2019

Abstract

We consider the problem of estimating the density of buyers and vendors in a nonlinear parabolic price formation model using measurements of the price and the transaction rate. Our approach is based on a work by Puel et al., see [20], and results in a optimal control problem. We analyse this problems and provide stability estimates for the controls as well as the unknown density in the presence of measurement errors. Our analytic findings are supported with numerical experiments.

1 Introduction

In this paper we use techniques developed in the field of data assimilation to predict the dynamics of a nonlinear parabolic free boundary price formation model proposed by Lasry & Lions in [16]. The Lasry-Lions (LL) model describes the price evolution of a single good traded between a large group of buyers and a large group of vendors. The price enters as a free boundary, at which trading takes place. After the realisation of a transaction, buyers and vendors immediately sell or rebuy the good at a shifted price. The shift in the price is due to the previously paid constant transaction costs. The situation detailed above can be described by the following nonlinear parabolic partial differential equation

$$\partial_t f - \frac{\sigma^2}{2}\partial_{xx} f = \Lambda(t)(\delta_{p(t)-a} - \delta_{p(t)+a}),\ x \in \Omega, t > 0, \tag{1.1a}$$

$$\Lambda(t) = -\frac{\sigma^2}{2}\partial_x f(p(t),t),\ f(p(t),t) = 0, \tag{1.1b}$$

$$f(x,0) = f_0(x),\ p(0) = p_0. \tag{1.1c}$$

The positive part $f^+ = \max(f,0)$ of the function $f = f(x,t)$ corresponds to the distribution of buyers over the price $x \in \Omega$, the negative part $f^{\cdot} = \min(f,0)$ to the is the vendor distribution over the price. The free boundary $p = p(t)$ corresponds to the price where $f(\cdot,t) = 0$, the function Λ to the total number of transactions executed at that price. The immediate placement and execution of new bids and orders after the trading event are modeled by the Delta Diracs at the shifted prices $p(t) + a$ and $p(t) - a$, where $a \in \mathbb{R}^+$ denotes the transaction

*Department Mathematik, Friedrich-Alexander Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen. e-mail:martin.burger@fau.de

†Technische Universität Chemnitz, Fakultät für Mathematik, Reichenhainer Str. 41, 09126 Chemnitz Germany. jfpietschmann@math.tu-chemnitz.de

‡Mathematics Institute, Warwick University, CV4 7AL Coventry, UK and Radon Institute for Computational and Applied Mathematics, Altenbergerstr. 69, 4040 Linz, Austria. e-mail: m.wolfram@warwick.ac.uk

costs. Random changes in the buyer and vendor distribution are included by a Laplacian with constant diffusivity $\sigma \in \mathbb{R}^+$. We assume that the initial distribution f_0 satisfies:

$$f_0(p_0) = 0,\ f_0(x) > 0 \text{ for } x < p_0 \text{ and } f_0(x) < 0 \text{ for } x > p_0, \text{ a.e. in } \Omega \tag{1.2}$$

and set w.l.o.g. $\frac{\sigma^2}{2} = 1$. System (1.1) can be posed on the positive real line $\Omega = \mathbb{R}^+$ or a bounded interval $\Omega = [0, x_{\max}]$, where $x_{\max}$ denotes the maximum price. We will consider (1.1) on the bounded interval only and impose homogeneous Neumann boundary conditions

$$\partial_x f = 0 \text{ on } \partial\Omega \tag{1.3}$$

to ensure that the total number of buyers and vendors is constant in time.
For convenience, we assume the initial price p_0 is normalized to 0 and only consider its relative change. Hence we work on the shifted domain $[-L, L]$, where $L = \frac{x_{\max}}{2}$. Altogether we will consider (1.1) with boundary condition (1.3) on $\Omega = [-L, L]$ throughout this manuscript.

The LL model (1.1) was analyzed in a series of papers, cf. [11, 19, 8, 4, 5]. Most available results are based on a nonlinear transformation of (1.1), which transforms the problem to the heat equation with nonlinear boundary conditions. This connection provides the main analytical ingredients to study existence and long time behavior of solutions to (1.1). Lasry and Lions introduced the model on the macroscopic level only, a more detailed microscopic interpretation of the trading process and the respective limit as the number of buyers and vendors tend to infinity was missing. This connection was established by Burger et al., who proved that the original LL model can be derived from a Boltzmann type model as the number of transactions tends to infinity, see [2]. In their approach trading events between buyers and vendors are modeled by "collisions", which can also be used to describe price dynamics in case of more general trading rules. The connection between the Boltzmann-type price formation model and the LL model (1.1) was further investigated in different asymptotic limits in [3]. The LL and Boltzmann-type price formation models are appealing in many respects, especially in terms of analytical tractability. However the resulting price process is deterministic and does not give any insights into connections between transactions rates, order flows or price volatility. Markowich et al., [18] considered a stochastic extensions of the original LL model. However this extension did not give realistic price dynamics either. Very recently Cont and Müller [10] proposed a stochastic partial differential equation with multiplicative noise, which reproduces statistical properties of real price dynamics.

In this paper we focus on the inverse problem of determining the buyer-vendor distribution given measurements of the price and the transaction rate on a time interval $[0, T]$. This distribution can then be used as an initial value and thus allows us to predict price and transaction rate for $t > T$. More specifically we will investigate the question

Problem I: Given measurements of the price $p(t)$ and the transaction rate $\Lambda(t)$ in some time interval $[0, T]$, is it possible to predict the price for times $t > T$?

Our approach is based on an optimal control approach proposed by J-P. Puel, see [20, 21]. It is based on a duality argument, which allows to reconsttuct the distribution f at the final time T. This is in contrast to standard data assimilation where one tries to recover the initial datum $f_0(x)$. We adapte the strategy of Puel et al. and use duality estimates to compute linear functionals of $f(T, x)$. These functionals involve the solution of optimal boundary control problems with PDE constraints. Optimal boundary control problems are well studied in the literature, see e.g. [17, 22, 13]. We will make use of an exact null controllability result for parabolic boundary control problems shown in [7]. Its proof is based on Carleman estimates, a technique commonly used to derive exact controllability results (and also uniqueness for inverse problems), see [23, 14] for details. A possible numerical realisation of Puel's strategy was presented in [9].

Our contributions to the subject of optimal control for parabolic free boundary problems and data assimilation in price formation models are the following:

- We present the first approach to reconstruct the buyer- and vendor distribution from measurements of price and transaction rate (to the author's knowledge).
- We generalise the data assimilation approach of Puel et al., see [20], to free boundary value problems and evolving domains.
- We provide stability estimates, which give novel insights into the influence of measurement errors on the price dynamics.
- We propose a computational strategy to implement the developed framework numerically.

This paper is organized as follows: The proposed framework is based on several analytic results, which will be presented in Section 2. The data assimilation problem itself is discussed in Section 3. Section 4 is devoted to stability in the presence of measurement errors and we conclude by presenting numerical experiments in Section 5.

2 Preliminary results

In this section we provide analytic tools and results of the forward problem and define the respective adjoint problem, which will be used in the optimal control approach.
The presented results rely on the following assumptions:

(A1) $f_0(p_0) = 0,\ f_0(x) > 0$ for $x < p_0$ and $f_0(x) < 0$ for $x > p_0$.

(A2) For every $t \in [0, T]$, there exists a constant $\underline{p} > a$ such that $-L + \underline{p} \leq p(t) \leq L - \underline{p}$.

Assumption (A1) is the necessary compatibility condition for the initial datum f_0 (which we already stated in (1.2)), while (A2) ensures that the price stays sufficiently far away from the interval boundaries. Note that the restriction on $p(t)$ is not severe in the context of inverse problems: Since we will assume later on that we know measurements of $p(t)$ in some time interval $[0, T]$, we can always chose the domain size L (within realistic bounds) such that the condition $p(t) \in (-L + a, L - a)$ is satisfied. As $p(t)$ is continuous, we also know it will stay in $(-L + a, L - a)$ for some time so that it is save to predict for $t > T$.

2.1 Nonlinear transformation of the model

We start by discussing the nonlinear transformation which converts (1.1) to a linear heat equation. This connection was exploited in almost all analytic results as well as computational methods. The idea is that the second derivative of f at $p(t)-a$ behaves like $\Lambda(t)\delta_{p(t)-a}$ while at $p(t)+a$ it behaves like $-\Lambda(t)\delta_{p(t)+a}$. Thus shifting the function by multiples of $\pm a$ and adding them up 'eliminates' the singularity on the right hand side. More precisely, for $\Omega = \mathbb{R}$, we define

$$F(x,t) = \begin{cases} \sum_{n=0}^{\infty} f^+(x+na,t), \ x < p(t) \\ -\sum_{n=0}^{\infty} f^-(x-na,t), \ x > p(t). \end{cases} \tag{2.1}$$

Then the function $F = F(x,t)$ satisfies the heat equation

$$\partial_t F(x,t) - \partial_{xx} F(x,t) = 0, \qquad x \in \mathbb{R}, t > 0, \tag{2.2a}$$

$$F(x,0) = F_0(x), \ x \in \mathbb{R}, \tag{2.2b}$$

with the transformed initial datum

$$F_0(x) = \begin{cases} \sum_{n=0}^{\infty} f_0^+(x+na), \ x < p(0) \\ -\sum_{n=0}^{\infty} f_0^-(x-na), \ x > p(0). \end{cases}$$

Since we consider (1.1) with homogeneous Neumann boundary conditions on a bounded domain, we only have a finite sum in (2.1) but obtain the following boundary conditions:

$$\partial_x F(-L,t) = \partial_x F(-L+a,t), \tag{2.3a}$$

$$\partial_x F(L,t) = \partial_x F(L-a,t), \tag{2.3b}$$

Note that the solution of the original LL model (1.1) can be computed by

$$f(x,t) = F(x,t) - F^+(x+a,t) + F^-(x-a,t).$$

2.2 Existence and regularity of the price

In the following we provide additional existence and regularity results for the direct problem. Note that these results are not optimal in terms of regularity. However, they are sufficient to define all quantities that we shall need in the sequel.

Theorem 2.1 (Existence of f, $p(t)$)**.** *Let $f_0 \in L^2(-L,L)$ and $p_0 \in (-L+a, L-a)$ satisfy (A1). Then the BVP (1.1) has a global solution conserving the total mass of buyers and vendors iff the zero level set $p = p(t)$ of the solution of (2.2)–(2.3) satisfies $p(t) \in (-L+a, L-a)$ for all $t > 0$. Then the free boundary $p(t)$ converges to the stationary price $p_\infty \in (-L+a, L-a)$.*

Proof. The proof is mainly based on the definition of the transformation (2.1), see [5] for details. □

Note that the stationary price is determined by the initial mass of buyers and vendors as well as the transaction rate a. In particular

$$p_\infty = \frac{2M^l L - a(M^l - M^r)}{2(M_l + M_r)} - \frac{L}{2} \tag{2.4}$$

where $M^l = \int_- L^{p_0} f_0(x)dx$ and $M^r = \int_{p_0}^{L} f_0(x)dx$. The presented analysis of the adjoint and assimilation problem relies on the following regularity result for the price $p = p(t)$.

Lemma 2.2 (Regularity of $p(t)$)**.** *Let $f_0 \in L^2(-L,L)$ and $p_0 \in (-L+a, L-a)$ satisfy (A1). Then $p(t) \in C^1([\varepsilon,T])$ for $\varepsilon > 0$.*

Proof. The results is a direct consequence of the fact that $F(x,t)$ is smooth in space and time for all $t > 0$ and of the boundedness of Λ. Indeed, differentiating the relation $F(p(t),t) = 0$ yields

$$p'(t) = \frac{\partial_t F(p(t),t)}{\Lambda(t)} = \frac{\partial_{xx} F(p(t),t)}{\Lambda(t)}, \tag{2.5}$$

and therefore

$$\sup_{t\in[\varepsilon,T]} p'(t) \leq \frac{\|\partial_{xx}F\|_{C([-L,L]\times[\varepsilon,T])}}{\underline{\Lambda}},$$

where the parabolic version of Hopf's Lemma applied at $x = p(t)$ ensures that $\underline{\Lambda} = \min_{t\in[\varepsilon,T]} \Lambda(t) > 0$. □

Remark 2.3. *The regularity of the price p as well as the buyer-vendor density f at the initial time is crucial to define the transformation between the time-dependent domains $[-L, p(t)]$ and $[L, p(t)]$ and the reference domain $[0,1]$ (see Subsection 2.3) but also for the exact controllability results of Theorem 3.3. Therefore we will work the temporal domain $[\varepsilon, T]$ instead of $[0,T]$ for some fixed $\varepsilon > 0$ in the following only.*

2.3 Evolving spaces and transformation to fixed domain

A crucial step in the subsequent analysis is the splitting of the domain Ω into the part left and right of the price $p(t)$ (illustrated in Figure 1). We introduce the domains

$$\Omega_\triangleleft = [-L, p(t)], \quad \Omega_\triangleright = [p(t), L], \text{ and } \Omega = [0,1],$$

as well as

$$Q_\triangleleft = \Omega_\triangleleft \times [\varepsilon, T], \quad Q_\triangleright = \Omega_\triangleright \times [\varepsilon, T], \text{ and } Q = \Omega \times [\varepsilon, T].$$

Following [1], we define evolving Bochner spaces on these domains. We present the construction for the left domain $\Omega_\triangleleft = [-L, p(t)]$ only, since the argument for the right domain is analogous. First denote by $H^1_\triangleleft(t) := H^1((-L, p(t))$ the evolving Hilbert space. Next we define the map $\phi_t : H^1_\triangleleft(\varepsilon) \to H^1_\triangleleft(t)$ by

$$\phi_t u(x) = u(\kappa x + \zeta L),$$

with $\kappa = \frac{p(\varepsilon)+L}{p(t)+L}$ and $\zeta = \frac{p(\varepsilon)-p(t)}{p(t)+L}$ for all $-L \leq x \leq p(t)$ and $\varepsilon \leq t \leq T$. The function ϕ_t is obviously continuous and reduces to the identity at $t = \varepsilon$. It is also a homeomorphism as its inverse

$$\phi_{-t} u(x) = u(\kappa^{-1} x + \zeta^{-1} L) \text{ for } -L \leq x \leq p(\varepsilon) \text{ and } \varepsilon \leq t \leq T,$$

is continuous as well. This allows us to introduce the evolving Bochner spaces (as in [1, Definition 2.7])

$$L^2_{H^1_\triangleleft} = \{u : [\varepsilon, T] \to \bigcup H^1_\triangleleft(t) \times \{t\}, t \mapsto (\bar{u}(t), t) \mid \phi_{-(\cdot)} \bar{u}(\cdot) \in L^2\left(\varepsilon, T; H^1_\triangleleft(\varepsilon)\right)\}, \tag{2.6}$$

$$L^2_{(H^1_\triangleleft)^*} = \left\{g : [\varepsilon, T] \to \bigcup (H^1_\triangleleft)^*(t) \times \{t\}, t \mapsto (\bar{g}(t), t) \mid \phi^*_{(\cdot)} \bar{g}(\cdot) \in L^2\left(\varepsilon, T; (H^1_\triangleleft(\varepsilon))^*\right)\right\}, \tag{2.7}$$

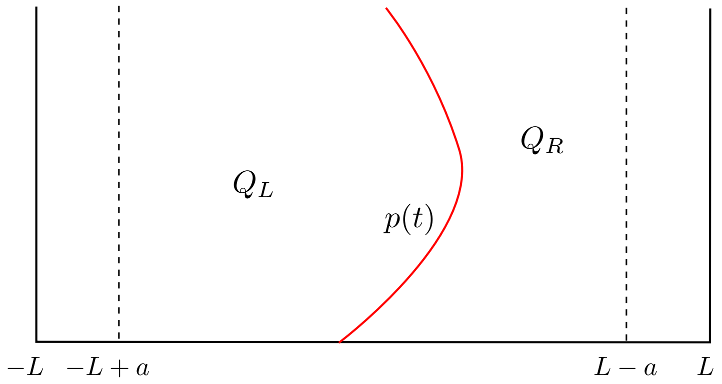


Figure 1: Sketch of the space-time domain and the splitting into $Q_\triangleleft$ (left part) and $Q_\triangleright$ (right part)

and, following again [1], make the identification of $u(t) = (\bar{u}(t), t)$ with $\bar{u}(t)$ for $u \in L^2_{H^1_\triangleleft}$ (and likewise in $L^2_{(H^1_\triangleleft)^*}$). The space of continuously differentiable functions on evolving Bochner spaces is given by

$$C^k_{H^1_\triangleleft} = \left\{\xi \in L^2_{H^1_\triangleleft} | \phi_{-(\cdot)}\xi(\cdot) \in C^k\left([\varepsilon, T]; H^1_\triangleleft(\varepsilon)\right)\right\} \qquad \text{for } k \in \{0, 1, \ldots\}.$$

Thus we can, as in [1, Definition 2.20], to give a notion of time (material) derivative as

$$\dot{\xi}(t) := \phi_t \left(\frac{d}{dt}\left(\phi_{-t}\xi(t)\right)\right) \in C^0_{H^1_\triangleleft},$$

for any $\xi \in C^1_{H^1_\triangleleft}$. Then we can finally define the space used for the notion of weak solutions, namely

$$W\left(H^1_\triangleleft, (H^1_\triangleleft)^*\right) = \left\{u \in L^2_{H^1_\triangleleft} \mid \dot{u} \in L^2_{(H^1_\triangleleft)^*}\right\}. \tag{2.8}$$

The definitions of the respective quantities $L^2_{H^1_\triangleright}$, $L^2_{(H^1_\triangleright)^*}$, $C^k_{H^1_\triangleright}$, and $W\left(H^1_\triangleright, (H^1_\triangleright)^*\right)$ are analogous.

While the previous definitions allow us to directly work in a noncylindrical domain, it is sometimes also useful consider the transformation to the fixed domain $Q = [0, 1] \times [\varepsilon, T]$. Hence we introduce transformations which map $Q_\triangleleft$ and $Q_\triangleright$ to Q:

$$T_\triangleleft : Q_\triangleleft \to Q, \qquad T_\triangleright : Q_\triangleright \to Q, \tag{2.9a}$$

$$(x, t) \mapsto \left(\frac{x + L}{L + p(t)}, t\right), \qquad (x, t) \mapsto \left(\frac{-L + x}{-L + p(t)}, t\right). \tag{2.9b}$$

Note that due to assumption (A1), $T_\triangleleft$ and $T_\triangleright$ are well-defined and that $T_\triangleright$ actually flips the domain, i.e. it swaps left and right boundary points.

2.4 Adjoint equations

The next ingredient will be two adjoint equations, posed on the domains $Q_\triangleleft$ and $Q_\triangleright$, respectively.

Definition 2.4 (Adjoint equations)**.** For any $\varepsilon > 0$, $\psi_\triangleleft \in L^2(-L, p(T))$, $\psi_\triangleright \in L^2(p(T), L)$, $u_\triangleleft$, $u_\triangleright \in L^2(\varepsilon, T)$ and $T > 0$, we introduce the backward in time adjoint equations

$$\begin{aligned} -\partial_t \Phi_\triangleleft(x,t) - \partial_{xx}\Phi_\triangleleft(x,t) &= 0, && \text{in } Q_\triangleleft && (2.10a)\\ \partial_x \Phi_\triangleleft(-L,t) &= 0, && \text{for } t \in [T,\varepsilon] && (2.10b)\\ \Phi_\triangleleft(p(t),t) &= u_\triangleleft(t), && \text{for } t \in [T,\varepsilon] && (2.10c)\\ \Phi_\triangleleft(x,T) &= \Psi_\triangleleft(x), && \text{for } x \in \Omega_\triangleleft. && (2.10d) \end{aligned}$$

and

$$\begin{aligned} -\partial_t \Phi_\triangleright(x,t) - \partial_{xx}\Phi_\triangleright(x,t) &= 0, && \text{in } Q_\triangleright && (2.11a)\\ \partial_x \Phi_\triangleright(L,t) &= 0, && \text{for } t \in [T,\varepsilon] && (2.11b)\\ \Phi_\triangleright(p(t),t) &= u_\triangleright(t), && \text{for } t \in [T,\varepsilon] && (2.11c)\\ \Phi_\triangleright(x,T) &= \Psi_\triangleright(x), && \text{for } x \in \Omega_\triangleright. && (2.11d) \end{aligned}$$

Applying the existence theory of, e.g. [1], for equations on evolving domains, we obtain the following theorem.

Theorem 2.5. *Let $p \in C^1([\varepsilon, T])$ be given. Then, for every $\Psi_\triangleleft \in L^2(\Omega_\triangleleft)$, $u_\triangleleft \in L^2(\varepsilon, T)$ and every $\Psi_\triangleright \in L^2(\Omega_\triangleright)$, $u_\triangleright \in L^2(\varepsilon, T)$ there exist unique solutions $\Phi_\triangleleft$ and $\Phi_\triangleright$ to* (2.10) *and* (2.11)*, respectively. Furthermore, we have*

$$\Phi_\triangleleft \in W\left(H^1_\triangleleft, (H^1_\triangleleft)^*\right), \qquad (2.12)$$

$$\Phi_\triangleright \in W\left(H^1_\triangleright, (H^1_\triangleright)^*\right). \qquad (2.13)$$

With the help of the transformations $T_\triangleleft$ and $T_\triangleright$, equation (2.10) and (2.11) can be transformed into a generic problem of the form

$$\begin{aligned} -\partial_t \Phi - a(t)\partial_{yy}\Phi + b(t) y \partial_y \Phi &= 0, && \text{for } (x,t) \text{ in } Q && (2.14a)\\ \partial_y \Phi(0,t) &= 0, && \text{for } t \in [T,\varepsilon] && (2.14b)\\ \Phi(1,t) &= u(t), && \text{for } t \in [T,\varepsilon] && (2.14c)\\ \Phi(y,T) &= \Psi(y), && y \in (0,1). && (2.14d) \end{aligned}$$

For (2.10) we define $(y,t) = T_\triangleleft(x,t)$ and compute

$$a(t) = \frac{1}{(p(t)+L)^2},\ b(t) = \frac{p'(t)}{(p(t)+L)},\ u(t) = u_\triangleleft(t) \text{ and } \Psi(y) = \Psi_\triangleleft((p(T)+L)y - L), \qquad (2.15)$$

while for (2.11) and $(y,t) = T_\triangleright$ we obtain

$$a(t) = \frac{1}{(p(t)-L)^2},\ b(t) = \frac{p'(t)}{(p(t)-L)},\ u(t) = u_\triangleright(t) \text{ and } \Psi(y) = \Psi_\triangleright((p(T)-L)y + L). \qquad (2.16)$$

Note that in view of Lemma 2.2 and Assumption (A1), the coefficients a and b are (in both cases) continuous and uniformly bounded by

$$\frac{1}{(2L - \underline{p})^2} < a(t) \le \frac{1}{\underline{p}^2} \quad \text{and} \quad 0 \le b(t) \le \frac{\|p\|_{C^1([0,T])}}{\underline{p}^2}, \qquad (2.17)$$

as there may be points with $p'(t) = 0$. Thus, standard existence and regularity results for linear diffusion–convection equations on fixed domains, such as [15, Theorem 5.2], can be used to ensure the solvability of (2.14).

3 Data assimilation problem

We now turn to the main part of this paper - the inverse or data assimilation problem I. In classic data assimilation approaches one would use the measurements of $p = p(t)$ and $\Lambda = \Lambda(t)$ on $[0, T]$ to reconstruct the initital datum $f_0(x)$ of (1.1). Here we follow an alternative approach proposed by Puel et al., see [20, 21], and estimate the buyer-vendor distribution at the final time, that is $f(x, T)$ instead. This requires the solution of additional optimal control problems, which are, however, well posed if an appropriate regularisation (penalty) is added.

To use Puel's strategy in our setting, we will estimate the densities of buyers and of vendors separately (that is on the right and left of the free boundary). The reconstruction is then based on the following two duality estimates:

Theorem 3.1. *Let $f_0 \in L^2(\Omega)$ satisfying assumption (A1) and let $f \in L^2(0, T; H^1(\Omega))$ be a solution to* (1.1) *with corresponding price, $p \in C^1([\varepsilon, T])$ satisfying (A2). Furthermore, let*

$$\Phi_\triangleleft \in W\left(H^1_\triangleleft, (H^1_\triangleleft)^*\right), \quad \Phi_\triangleright \in W\left(H^1_\triangleright, (H^1_\triangleright)^*\right) \tag{3.1}$$

satisfy (2.10) *and* (2.11)*, respectively. Then, the following duality estimates*

$$\int_{-L}^{p(T)} f(x,T)\Psi_\triangleleft(x)\,dx = \int_{-L}^{p(\varepsilon)} f(x,\varepsilon)\Phi_\triangleleft(x,\varepsilon)\,dx + \int_\varepsilon^T \Lambda(t)(\Phi_\triangleleft(p(t)-a) - u_\triangleleft(t))\,dt, \tag{3.2a}$$

$$\int_{p(T)}^{L} f(x,T)\Psi_\triangleright(x)\,dx = \int_{p(\varepsilon)}^{L} f(x,\varepsilon)\Phi_\triangleright(x,\varepsilon)\,dx + \int_\varepsilon^T \Lambda(t)(u_\triangleright(t) - \Phi_\triangleright(p(t)+a))\,dt, \tag{3.2b}$$

hold for arbitrary functions $u_\triangleleft$, $u_\triangleright \in L^2(0, T)$ and every $\varepsilon > 0$.

Proof. We prove the first estimate only, since the argument for (3.2b) is the same. We have

$$\begin{aligned}
\int_{-L}^{p(T)} f(x,T)\Psi_\triangleleft(x)\,dx - \int_{-L}^{p(\varepsilon)} f(x,\varepsilon)\Phi_\triangleleft(x,\varepsilon)\,dx &= \int_\varepsilon^T \int_{-L}^{p(t)} \partial_t(f(x,t)\Phi_\triangleleft(x,t))\,dxdt \\
&= \int_\varepsilon^T \int_{-L}^{p(t)} [\partial_{xx} f(x,t) + \lambda(t)\delta_{p(t)-a}]\Phi_\triangleleft(x,t) - f(x,t)\partial_{xx}\Phi_\triangleleft(x,t)\,dxdt \\
&= \int_\varepsilon^T (\partial_x f\Phi_\triangleleft)|_{x=-L}^{x=p(t)} - (f\partial_x\Phi_\triangleleft)|_{x=-L}^{x=p(t)}\,dt + \int_\varepsilon^T \Lambda(t)\Phi_\triangleleft(p(t)-a,t)\,dt \\
&= \int_\varepsilon^T \Lambda(t)(u_\triangleleft(t) + \Phi_\triangleleft(p(t)-a,t))\,dt.
\end{aligned}$$

where we have used the boundary condition (1.3), $f(p(t), t) = 0$ and the definition of Λ. □

Now we will use (3.2a)–(3.2b) to determine $f(x, T)$. Since the choice of $\Psi_\triangleleft$ and $\Psi_\triangleright$ in (3.2a) and (3.2b) is arbitrary and the last term on the right hand side contains only known (i.e. computed or measured) quantities, we could obtain a linear functional of $f(x, T)$. The only unknowns are the first terms on the respective right hand sides. But since we are free to choose arbitrary boundary data $u_\triangleleft$ and $u_\triangleright$, this leads to the null–controllability problems for (2.10)–(2.11). Indeed, if we can chose $u_\triangleleft$ and $u_\triangleright$ such that $\Phi_\triangleleft(x, \varepsilon) = 0$ and $\Phi_\triangleright(x, \varepsilon) = 0$, the unknown terms in both orthogonality relations drop out and we can reconstruct $f(x, T)$.

3.1 Optimal control problem

To conduct the strategy outlined above, we have to solve the optimal control problems

$$\min_{u_\triangleleft \in L^2(\varepsilon,T)} \frac{1}{2}\int_{-L}^{p(\varepsilon)} \Phi_\triangleleft(x,\varepsilon)^2 dx \quad \text{subject to (2.10)}, \tag{3.3}$$

$$\min_{u_\triangleright \in L^2(\varepsilon,T)} \frac{1}{2}\int_{p(\varepsilon)}^{L} \Phi_\triangleright(x,\varepsilon)^2 dx \quad \text{subject to (2.11)}. \tag{3.4}$$

Since the structure of both problems is the same, we will only discuss the first one. To increase readability, we will drop the subscript $\triangleleft$ and write u, ϕ, ... instead of $u_\triangleleft$, $\phi_\triangleleft$ from now on. The next result states that the optimal control problem is indeed exactly null-controllable in the sense of the following definition.

Definition 3.2. We say that problem (3.3) is *exactly null-controllable*, if for every initial datum $\Psi \in L^2(\Omega_\triangleleft)$ to (2.10) there exists $\bar{u} \in L^2(\varepsilon, T)$ such that the solution Φ to (3.3) with control $u = \bar{u}$ satisfies $\Phi(x,\varepsilon) = 0$.

The following exact boundary controllability result is based on [7, Theorem 2.3], slightly extended and adapted to our situation. The theorem reads as follows.

Theorem 3.3 (Exact null-controllability)**.** *For every $\Psi \in L^2(\Omega_\triangleleft)$, there exists at least one control $u \in L^2(\varepsilon,T)$ such that the solutions Φ of (2.10) satisfies $\Phi(x,\varepsilon) = 0$ on $\Omega_\triangleleft$. Furthermore, there exists a constant C which depends on $p(t)$, L and T such that*

$$\|\bar{u}\|_{L^2(\varepsilon,T)} \le C\|\Psi\|_{L^2(\Omega_\triangleleft)} \tag{3.5}$$

holds with $\bar{u}$ being the control of minimum L^2–Norm.

Proof. The regularity of the price p allows us to transform the problem to a fixed domain using $T_\triangleleft$ defined in (2.9). Hence we only consider equations of type (2.14). First we observe that for any positive $\delta > 0$, any solution Φ to (2.14) with initial datum Ψ is, by standard parabolic regularity [12, Chapter 7.1], in $L^2(\varepsilon+\delta, T); H^1(0,1))$ with the estimate

$$\|\Phi\|_{L^2(\varepsilon+\delta,H^1(0,1))} \le C\|\Psi\|_{L^2(0,1)}. \tag{3.6}$$

Thus, w.l.o.g. we can assume that already $\Psi \in H^1(0,1)$ holds. Since by lemma 2.2, $p \in C^1([\varepsilon,T])$ (and thus the coefficients a and b in (2.14) are continuous) we can apply [7, Theorem 2.3] to conclude the requested boundary controllability. The continuity estimate (3.5) then follows by combining (3.6), the respective estimate from [7, Theorem 2.1] for the distributed control problem and a standard trace inequality.

□

In order to be able to numerically solve the optimal control problem, we introduce the following regularized version

$$\min_{u\in L^2(\varepsilon,T)} \frac{1}{2}\int_{-L}^{p(\varepsilon)} \Phi(x,\varepsilon)^2 dx + \frac{\alpha}{2}\int_{\varepsilon}^{T} u(t)^2\, dt \text{ subject to (2.10)} \tag{3.7}$$

Standard arguments guarantee the existence of a unique minimizer, see e.g. [22, Section 3.5]. Calculating the derivatives of the corresponding Lagrange functional

$$\begin{aligned}\mathcal{L} = \frac{1}{2}\int_{-L}^{p(\varepsilon)} \Phi(x,\varepsilon)^2 dx + \frac{\alpha}{2}\int_{\varepsilon}^{T} u(t)^2\, dt \\ + \int_{\varepsilon}^{T}\int_{-L}^{p(t)} G(x,t)\left[-\partial_t\Phi(x,t) - \partial_{xx}\Phi(x,t)\right]\, dxdt,\end{aligned} \tag{3.8}$$

we obtain the first order optimality system

$$\partial_t G(x,t) - \partial_{xx} G(x,t) = 0, \qquad \text{in } Q_\triangleleft \tag{3.9a}$$

$$\partial_x G(-L,t) = 0, \qquad \text{for } t > \varepsilon \tag{3.9b}$$

$$G(p(t),t) = 0, \qquad \text{for } t > \varepsilon \tag{3.9c}$$

$$G(x,\varepsilon) = -\Phi(x,\varepsilon), \quad \text{in } \Omega, \tag{3.9d}$$

where Φ satisfies the adjoint equation (2.10) and the coupling

$$\alpha u(t) + \partial_x G(p(t),t) = 0, \quad \text{for } t > \varepsilon. \tag{3.10}$$

The following results examine the convergence of u as $\alpha \to 0$. The proofs are using the same techniques as in [21], yet adapted to our boundary control problem.

Theorem 3.4. *For every $\alpha > 0$, denote by (u_α, ϕ_α) the corresponding solution to* (3.7)*. Then we have*

$$u_\alpha \to \bar{u} \text{ in } L^2(\varepsilon,T) \qquad \text{as } \alpha \to 0, \tag{3.11}$$

$$\Phi_\alpha \to \bar{\Phi} \text{ in } C([\varepsilon,T]; H^1(\Omega_\triangleleft)) \qquad \text{as } \alpha \to 0, \tag{3.12}$$

where $\bar{u}$ is the solution to the optimal control problem (3.3) *having minimal L^2-norm and Φ_α and $\bar{\Phi}$ are the solutions to* (2.10) *with boundary data u_α and $\bar{u}$, respectively.*

Proof. By Theorem 3.3, we know that there exists at least one function solving the exact null controllability problem. Thus, the set of all these controls in $L^2(\varepsilon,T)$ is nonempty. As it is also convex and closed, there exists a unique $\bar{u}$ having minimal L^2-norm. Since u_α minimizes the functional (3.7) among all function in $L^2(\varepsilon,T)$ we have

$$\frac{1}{2}\int_{-L}^{p(\varepsilon)} \Phi_\alpha(x,\varepsilon)^2 dx + \frac{\alpha}{2}\int_{\varepsilon}^{T} u_\alpha(t)^2\, dt \le \frac{\alpha}{2}\int_{\varepsilon}^{T} \bar{u}(t)^2\, dt \tag{3.13}$$

which implies the (uniform in α) bound

$$\frac{1}{2}\int_{\varepsilon}^{T} u_\alpha(t)^2\, dt \le C. \tag{3.14}$$

Thus, we can extract a subsequence, again labeled u_α that converges weakly to some $\tilde{u}$ in $L^2(\varepsilon,T)$. Using the weak formulation of (2.10) and an Aubin-Lions argument, we see that this is sufficient to obtain the convergence

$$\Phi_\alpha \to \tilde{\Phi} \text{ in } C([\varepsilon,T]; H^1(-L,p(t))), \qquad \text{as } \alpha \to 0$$

and (3.13) implies $\tilde{\Phi}(\varepsilon, x) = 0$. Thus, arguing as in the proof of [21, Theorem 2.12], we can use the fact that $\bar{u}$ has minimal norm as well as the lower semicontinuity of the norm w.r.t weak convergence to obtain that $\tilde{u} = \bar{u}$. This argument also implies norm convergence and the uniqueness of the limit then finally yields

$$u_\alpha \to \bar{u} \text{ in } L^2(\varepsilon, T).$$

This also implies $\tilde{\Phi} = \Phi$ which completes the proof. □

Remark 3.5. *Understanding the optimal control problem* (3.3) *(or* (3.4)*) as Tikhonov regularisation, one could ask for convergence rates of u_α to $\bar{u}$ as $\alpha \to \infty$. Indeed, such rates could be expected under appropriate source conditions on $\bar{u}$. The interesting point now is to understand the influence of $p(t)$ in the definition of the forward operator in the characterisation of such conditions and also how perturbation in p would influence them. We leave this question for future research.*

4 Stability in the presence of measurement errors

Assume we have measurements of two different prices p_1 and p_2 as well as two different transaction rates $\Lambda_1(t)$ and $\Lambda_2(t)$. Can we control the difference in the reconstructions $f_1(x,T)$ and $f_2(x,T)$ as well as the future predicted prices $p_1(t)$ and $p_2(t)$ for $t > T$ in terms of these differences? In this section we will give a positive answer to this question based on the following strategy

1. Estimate the error in the optimal controls u_1 and u_2 in terms of the error in p_1 and p_2 (Lemma 4.2).
2. Estimate the error in the respective reconstructions $f_1(x,T)$ and $f_2(x,T)$ in terms of errors in price and transaction rate (Lemma 4.3).
3. Use these results to predict errors in the future price (Lemma 4.7).

Note however that for the last point we need to make additional regularity assumptions on the reconstructed final data that do not directly follow from our analysis (see Remark 4.5 for details). We start by assuming

(A3) $\|p_1 - p_2\|_{C^1([\varepsilon,T])} \leq \delta_p$ and $\|\Lambda_1 - \Lambda_2\|_{L^2((\varepsilon,T))} \leq \delta_\lambda$.

W.l.o.g. we only consider the optimality system related to (3.3), i.e. the left part $\Omega_\triangleleft = [-L, p(t)]$ and again drop the subscript $\triangleleft$. Moreover, we transform all equations to the unit interval $[0,1]$, so that the optimality system reads as

$$\begin{aligned} -\partial_t \Phi - a(t)\partial_{yy}\Phi + b(t)y\partial_y\Phi &= 0, && \text{in } Q && (4.1a)\\ \partial_y \Phi(0,t) &= 0, && \text{for } t > \varepsilon && (4.1b)\\ \Phi(1,t) &= u(t), && \text{for } t > \varepsilon && (4.1c)\\ \Phi(y,T) &= \Psi(y), && \text{for all } 0 \leq y \leq 1. && (4.1d)\end{aligned}$$

$$
\begin{aligned}
\partial_t G - a(t)\partial_{yy} G - b(t) y \partial_y G &= 0, && \text{in } Q && (4.1e)\\
\partial_y G(0,t) &= 0, && \text{for } t > \varepsilon && (4.1f)\\
G(1,t) &= 0, && \text{for } t > \varepsilon && (4.1g)\\
G(y,\varepsilon) &= -\Phi(y,\varepsilon), && \text{for all } 0 \le y \le 1. && (4.1h)
\end{aligned}
$$

and the coupling condition

$$
\alpha u(t) + \frac{1}{(p(t)+L)} \partial_y G(1,t) = 0, \quad \text{for } t > \varepsilon, \tag{4.1i}
$$

with $a(t)$ and $b(t)$ as defined in (2.15). Note that the transformed primal and dual equations are still adjoint to one another, yet now with respect to the scalar product

$$
(u,v) := \int_\varepsilon^T \int_0^1 (p(t)+L) uv \, dx dt. \tag{4.2}
$$

Lemma 4.1. *Let Φ and G be the solutions to* (4.1a) *and* (4.1e), *respectively. Then we have*

$$
\begin{aligned}
\|\Phi\|_{L^\infty((T,\varepsilon);L^2(0,1))} + \|\sqrt{a}\Phi\|_{L^2((T,\varepsilon);H^1(0,1))} &\le C_1 \left(1 + \|\Psi\|_{L^2((0,1))}\right),\\
\|G\|_{L^\infty((\varepsilon,T);L^2(0,1))} + \|\sqrt{a}G\|_{L^2((\varepsilon,T);H^1(0,1))} &\le C_2 \|\Phi(\cdot,\varepsilon)\|_{L^2((0,1))},
\end{aligned}
$$

with $C_1 = C_1(\alpha, \underline{p}, L, T)$ and $C_2 = C_2(\underline{p}, L, T)$.

Proof. These are standard estimates obtained choosing Φ and G as test functions in the weak formulation of (4.1a) and (4.1e), respectively. For the first estimate, we additionally used the L^2-bound (3.14) on the boundary control, which introduced the α-dependence in C_1. □

Now we are able to prove stability of the optimal control problem in terms of measurement errors in the price.

Lemma 4.2 (Stability of u)**.** *Consider two different prices $p_1(t)$ and $p_2(t)$ such that $p_1(\varepsilon) = p_2(\varepsilon)$ and $\|p_1 - p_2\|_{C^1([\varepsilon,T])} \le \delta_p$. Denote by Φ_1 and Φ_2 and G_1 and G_2 the solutions to* (4.1a) *and* (4.1e) *with $p = p_1$ and $p = p_2$, respectively. Then the following stability estimate for the controls u_1 and u_2 holds:*

$$
\int_0^1 (\Phi_1(x,\varepsilon) - \Phi_2(x,\varepsilon))^2 \, dx + \frac{\alpha}{2} \int_\varepsilon^T (u_1(t) - u_2(t))^2 \, dt \le C_3(\alpha, \underline{p}, L, T, \Psi)\delta_p^2.
$$

Proof. For each p_i (and corresponding a_i, b_i), we denote by G_i, Φ_i and u_i the corresponding solutions to the optimality system (4.1a)–(4.1i) and furthermore

$$
\bar{\Phi} = \Phi_1 - \Phi_2, \quad \bar{G} = G_1 - G_2.
$$

Then, $\bar{\Phi}$ and $\bar{G}$ satisfy, in the weak sense, the equations

$$
\begin{aligned}
-\partial_t \bar{\Phi} - a_1(t)\partial_{yy}\bar{\Phi} + b_1(t) y \partial_y \bar{\Phi} &= -(a_1 - a_2)\partial_{yy}\Phi_2 + (b_1 - b_2) y \partial_y \Phi_2, \text{in } Q && (4.3a)\\
\partial_y \bar{\Phi}(0,t) &= 0, && \text{for } t > \varepsilon \quad (4.3b)\\
\bar{\Phi}(1,t) &= u_1(t) - u_2(t), && \text{for } t > \varepsilon \quad (4.3c)\\
\bar{\Phi}(y,T) &= 0, && 0 \le y \le 1. \quad (4.3d)
\end{aligned}
$$

and

$$\begin{aligned}
\partial_t \bar{G} - a_1(t)\partial_{yy}\bar{G} - b_1(t)y\partial_y\bar{G} &= -(a_1-a_2)\partial_{yy}G_2 - (b_1-b_2)y\partial_y G_2, \text{ in } Q && (4.3e)\\
\partial_y G(0,t) &= 0, \qquad \text{for } t>\varepsilon && (4.3f)\\
G(1,t) &= 0, \qquad \text{for } t>\varepsilon && (4.3g)\\
G(y,\varepsilon) &= -\bar{\Phi}(y), \qquad 0\le y\le 1. && (4.3h)
\end{aligned}$$

Note that the following calculations are formal since for now we only know existence of weak solutions and therefore some of the integrals are not defined. In the end we arrive, however, at an estimate which is again well defined and could can be obtained rigorously by directly working with weak solutions. We chose this way of presentation as we believe it to be easier to follow. Thus (formally) taking equation (4.3a) and testing it with $\bar{G}$ (with respect to the scalar product (4.2)) yields

$$\begin{aligned}
&\int_Q (p(t)+L)\bar{G}[-\partial_t\bar{\Phi} - a_1(t)\partial_{yy}\bar{\Phi} + b_1(t)y\partial_y\bar{\Phi}]\,dxdt\\
&\qquad = \int_Q (p(t)+L)\bar{G}[-(a_1-a_2)\partial_{yy}\Phi_2 + (b_1-b_2)y\partial_y\Phi_2]\,dxdt
\end{aligned}$$

Integrating by parts on the left hand side, using (4.3e) and the boundary conditions results in

$$\begin{aligned}
&(p(t)+L)\int_0^1 \bar{\Phi}(x,\varepsilon)^2\,dx + \alpha\int_\varepsilon^T a_1(t)(u_1(t)-u_2(t))^2(p(t)+L)\,dt\\
&\quad = \int_Q (p(t)+L)\left\{[-(a_1-a_2)[\partial_{yy}\Phi_2\bar{G} - \partial_{yy}G_2\bar{\Phi}] + (b_1-b_2)y[\partial_y\Phi_2\bar{G} - \partial_y G_2\bar{\Phi}]\right\}\,dxdt.
\end{aligned}$$

A final integration by parts to remove the second derivatives on the right hand side gives

$$\begin{aligned}
&(p(t)+L)\int_0^1 \bar{\Phi}(x,\varepsilon)^2\,dx + \alpha\int_\varepsilon^T a_1(t)(u_1(t)-u_2(t))^2(p(t)+L)\,dt\\
&\quad = \int_Q (p(t)+L)\left\{[-(a_1-a_2)[-\partial_y\Phi_2\partial_y G_1 + \partial_y G_2\partial_y\Phi_1] + (b_1-b_2)y[\partial_y\Phi_2\bar{G} - \partial_y G_2\bar{\Phi}]\right\}\,dxdt\\
&\qquad + \alpha\int_\varepsilon^T (a_1-a_2)(u_1-u_2)u_2(p(t)+L)\,dt.
\end{aligned}$$

Using the estimates of Lemma 4.1, the boundedness of u in L^2 (see (3.14)) and Cauchy's inequality applied to the last term on the right hand side, we have

$$\int_0^1 \bar{\Phi}(x,\varepsilon)^2\,dx + \frac{\alpha}{2}\int_\varepsilon^T (u_1(t)-u_2(t))^2\,dt \le C_4(\underline{p},L,\Psi,\alpha)(\|a_1-a_2\|^2_{L^\infty(\varepsilon,T)} + \|b_1-b_2\|^2_{L^\infty(\varepsilon,T)}),$$

where we also used the lower bounds (2.17) on a and Assumption (A3) to estimate the expression $(p(t)+L)$ from below by $\underline{p}$ and above by $L-\underline{p}$. Using again (2.17) yields

$$\|a_1-a_2\|_{L^\infty((\varepsilon,T))} \le C_4(\underline{p},L)\delta_p, \text{ and } \|b_1-b_2\|_{L^\infty((\varepsilon,T))} \le C_5(\underline{p},L)\delta_p.$$

Combining this with the previous estimate yields the assertion. □

For the second step of our strategy, we return to the orthogonality relation (3.2a) which, transformed to $[0,1]$, reads as

$$(p(T)+L)\int_0^1 f(x,T)\Psi(x)\,dx = (p(\varepsilon)+L)\int_0^1 f(x,\varepsilon)\Phi(x,\varepsilon)\,dx \\ + \int_\varepsilon^T \Lambda(t)(\Phi(p(t)-a)-u(t))\,dt. \tag{4.4}$$

In the presence of errors in p and Λ we obtain two different relations and the following stability result. Note that the above results on the adjoint equations imply solvability for Φ with continuous dependence on the initial value for any $\Psi \in L^2([0,1])$. Hence, the duality relation uniquely defines $f(\cdot,T) \in L^2([0,1])$ when given $f(\cdot,\epsilon) \in L^2([0,1])$. There is further stable dependence of $f(\cdot,T)$ on the errors in the price and transaction rates, which we make precise by the following result:

Lemma 4.3 (Stability of $f(x,T)$)**.** *Let p_1, p_2 and Λ_1, Λ_2 be given functions which satisfy Assumption (A3) and denote by $f_1(x,T)$ and $f_2(x,T)$ the corresponding reconstructed prices calculated using* (4.4)*. Then we have*

$$\int_0^1 (f_1(x,T)-f_2(x,T))^2\,dx \le C_6(\underline{p}, f(\cdot,\varepsilon), \Psi, \alpha, L, T)(\delta_p+\delta_\lambda).$$

Proof. Subtracting (4.4) for (p_1,λ_1,u_1) and (p_2,λ_2,u_2) yields

$$\begin{aligned}
&(p_1(T)+L)\int_0^1 (f_1(x,T)-f_2(x,T))\Psi(x)\,dx = \\
&= \underbrace{(p_1(\varepsilon)+L)\int_0^1 f(x,\varepsilon)(\Phi_1(x,\varepsilon)-\Phi_2(x,\varepsilon))\,dx + (p_1(\varepsilon)-p_2(\varepsilon))\int_0^1 f(x,\varepsilon)\Phi_2(x,\varepsilon)\,dt}_{=:(I)} \\
&\quad + \underbrace{\int_\varepsilon^T \Lambda_1(t)[\Phi_1(p_1(t)-a)-u_1(t))-\Phi_2(p_2(t)-a)-u_2(t))]\,dt}_{=:(II)} \\
&\quad + \underbrace{\int_\varepsilon^T [\Lambda_1(t)-\Lambda_2(t)](\Phi_2(p_1(t)-a)-u_2(t))\,dt}_{=:(III)} - \underbrace{(p_1(T)-p_2(T))\int_0^1 f_2(x,T)\Psi(x)\,dx}_{=:(IV)}
\end{aligned}$$

We estimate each term of the right hand side separately

$$\begin{aligned}
(I) &\le (L-\underline{p})\|f(\cdot,\varepsilon)\|_{L^2((0,1))}\|\Phi_1(\cdot,\varepsilon)-\Phi_2(\cdot,\varepsilon)\|_{L^2(0,1)} \\
&\quad + \|p_1-p_1\|_{L^\infty(\varepsilon,T)}\|f(\cdot,\varepsilon)\|_{L^2((0,1))}\|\Phi_2(\cdot,\varepsilon)\|_{L^2((0,1))} \\
&\le C_7(\underline{p}, f, \Psi, \alpha, L, T)\delta_p,
\end{aligned}$$

where we used Lemmata 4.1 and 4.2. Next we have

$$\begin{aligned}
(II) &\le \|\lambda_1\|_{L^2(\varepsilon,T)}\|u_1-u_2\|_{L^2(\varepsilon,T)} \\
&\quad + \int_\varepsilon^T \Lambda_1[\phi_1(p_1(t)-a)-\phi_1(p_2(t)-a)+\phi_1(p_2(t)-a)-\phi_2(p_2(t)-a)]\,dt \\
&\le C_8\|p_1(t)-p_2(t)\|_{L^\infty(\varepsilon,T)} + C_9\|\Phi_1-\Phi_2\|_{L^2((0,1))} \le C_{10}\delta_p
\end{aligned}$$

using that for positive times $t \geq \varepsilon$ (and away from the boundary) ϕ_1 is Lipschitz continuous. Next we have

$$(III) \leq C_{11}\|\Lambda_1 - \Lambda_2\|_{L^2(\varepsilon,T)} \leq C_{11}\delta_\lambda.$$

and finally

$$(IV) \leq C_{12}\|p_1 - p_2\|_{C^1(\varepsilon,T)} \leq C_{12}\delta_p.$$

Combining all estimates and taking the supremum over all $\Psi \in L^2((0,1))$ with $\|\Psi\|_{L^2((0,1))} = 1$, we finally obtain

$$\|f_1(x,T) - f_2(x,T)\|_{L^2((0,1))} \leq C_{13}\delta_p + C_{11}\delta_\lambda \tag{4.5}$$

Taking $C_6 = \max(C_{13}, C_{11})$ yields the assertion. □

Remark 4.4. *The estimates of Lemma 4.2 and 4.3 show that, for $\alpha > 0$, the recontruction of the unkown buyer vendor distribution $f(x,T)$ is acutally a well-posed problem, at least for suffiently smooth perturbations of p. This is due to the fact that we are solving a regularized optimization problem. The price to pay is that the term involving $f(x,\varepsilon)$ in* (4.4) *does not vanish. However, since $f(x,\varepsilon)$ is fixed, is does not appear in our stability estimates.*

For the next result, we choose perturbed prices p_1 and p_2 such that $|p_1(T) - p_2(T)| < 2a$ and assume w.l.o.g. that $p_1(T) \leq p_2(T)$ and make the following additional assumptions:

(A4) $\delta_p < a$,

(A5) $f_1(x,T),\ f_2(x,T) \in L^2(-L,L) \cap H^4(I)$ with $I \subset (p_2(T) - a, p_1(T) + a)$

(A6) $\|f_1(x,T) - f_2(x,T)\|_{H^4(I)} \leq C_6(\delta_p + \delta_\lambda)$

Remark 4.5. *We mention that indeed it is natural to assume strong regularity of f in a neighbourhood of $p(T)$ for $T > 0$, since it locally arises as the solution of a heat equation. On the other hand we need to expect some singularities around $p(T) - a$ and $p(T) + a$ due to the singular source terms. Thus (A5) seems completely natural for forwards solutions of the price formation model. Moreover, it can also be verified that $f(\cdot,T)$ reconstructed via* (4.4) *has local H^4-regularity, which follows from using Ψ supported in I and an analysis of the solution of the parabolic equation for Φ, which can be estimated in terms of the H^{-4} norm of the initial value.*

In the following we analyze the forward propagation for $t > T$ in a small time interval. We denote the new initial value by $f_{i,0} := f_i(\cdot,T)$. First note that using the same localisation strategy as in [19] (i.e. multiplying the solution to (1.1) with a smooth cut-off function that has support inside the interval I), implies

$$\begin{aligned}\|f_i\|_{L^\infty\left((T,T+\gamma);H^\beta(I_2)\right)} &\leq C_{14}\left(\|f_i\|_{L^2(((T,T+\gamma));H^1((-L,L)))} + \|f_{i,0}\|_{H^\beta(I)}\right) \\ &\leq C_{15}\left(\|f_{i,0}\|_{L^2((-L,L))} + \|f_{i,0}\|_{H^\beta(I)}\right) \text{ for } \beta \leq 4,\end{aligned} \tag{4.6}$$

with $\gamma > 0$ to be fixed later on and where f_i is the solution to (1.1) with the reconstructed initial datum $f_i(x,T)$ that additionally satisfies (A4)-(A6). Furthermore, I_2 is an interval that is compactly contained in I. This allows us to derive the following estimates on terms of the form $(\partial_x K_N) * f$, where we denote by $K_N(x,t)$ the heat kernel with Neumann boundary conditions on $[-L,L]$, see e.g. [6, Section 6.4], and furthermore use the notation

$$K_N^T(x,t) := K_N(x,t-T).$$

Lemma 4.6. *For given $T>0$, initial values f_0, $f_{1,0}$ and $f_{2,0}$ at time T satisfying (A4), (A5) we have for $t \in [T, T+\gamma]$ with γ sufficiently small*

$$|((\partial_x K_N^T) * (f_{1,0} - f_{2,0}))(p(t),t)| \le C_{16} \left(\|f_{1,0} - f_{2,0}\|_{L^2((-L,L))} + \|f_{1,0} - f_{2,0}\|_{H^4(I)} \right).$$

For two continuous functions $p_1(t)$ and $p_2(t)$, we have

$$|((\partial_x K_N^T) * f_0)(p_1(t),t) - ((\partial_x K_N) * f_0)(p_2(t),t)| \le C_{17} \|p_1 - p_2\|_{C([0,T])}.$$

Proof. First note that $(\partial_x K_N^T) * f$ is the solution to the heat equation with homegeneous Neumann boundary condition, zero right hand side and initial datum f. Then, the first estimate is a direct consequence of (4.6) applied to such an solution with initial datum $f_{1,0} - f_{2,0}$. The second one follows from the fact that, as for t sufficiently small, $p_1(t)$ and $p_2(t)$ are in I_2 and thus, using again (4.6), the derivative of a solution to the heat equation that appears on the left hand side is Lipschitz continuous. □

We are now in a position to state the stability result for future prices.

Lemma 4.7. *Let assumptions (A3)–(A6) be satisfied and denote by f_1 and f_2 the solution to* (1.1) *on the time interval $[T, T+\gamma]$ with initial data $f_1(x,T)$ and $f_2(x,T)$, reconstructed from measurements p_1, Λ_1 and p_2, Λ_2 in $[0,T]$, respectively. Then there exists a constant $\gamma > 0$ and we the corresponding prices p_1 and p_2 for $t \in (T, T+\gamma)$ satisfy the estimate*

$$\|p_1(t) - p_2(t)\| \le C_{18} e^{C_{19}(t-T)} (\delta_p + \delta_\Lambda).$$

Proof. Due to assumption (A5) we can invoke [19][Lemma 2.5] to show that for γ sufficiently small (depending on $f_i(x,T)$, $i=1,2$) the corresponding transaction rates Λ_1, Λ_2 are strictly positive on $[T, T+\gamma]$. Furthermore, (A5) implies that $p_1(t)$, $p_2(t)$ are in $C^1([T, T+\gamma])$. Now Duhamel's formula allows us to express the solutions $f_i(x,t)$ to (1.1) as

$$f_i(x,t) = K_N^T * f_{i,0} + \int_T^t \Lambda_i(\tau) [K_N^T(x - p_i(\tau) + a, t-\tau) - K_N^T(x - p_i(\tau) - a, t - \tau)] \, d\tau \tag{4.7}$$

Taking the space derivative and evaluating at $x = p_i(t)$ we obtain

$$\begin{aligned} \Lambda_i(t) = \partial_x f_i(p(t),t) &= (\partial_x K_N^T) * f_{i,0}(p(t),t) \\ &+ \int_T^t \Lambda_i(\tau) [\partial_x K_N^T(p_i(t) - p_i(\tau) + a, t - \tau) - \partial_x K_N^T(p_i(t) - p_i(\tau) - a, t - \tau)] \, d\tau. \end{aligned} \tag{4.8}$$

Subtracting (4.8) for $i = 1,2$ and using the linearity of the convolution, we obtain

$$\begin{aligned} \Lambda_1(t) - \Lambda_2(t) = (\partial_x K_N^T) &* (f_{2,0}(p_2(t),t) - (\partial_x K_N^T * (f_{2,0}(p_1(t),t) \\ &+ (\partial_x K_N^T) * (f_{2,0} - f_{1,0})(p_1(t),t) \\ &+ \int_0^t (\Lambda_1(\tau) - \Lambda_2(\tau)) \, \theta_1(t,\tau) + \Lambda_2(\tau) \left(\theta_1(t,\tau) - \theta_2(t,\tau) \right) \, d\tau, \end{aligned} \tag{4.9}$$

with $\theta_i(t,\tau) = [\partial_x K_N^T(p_i(t) - p_i(\tau) + a, t-\tau) - \partial_x K_N(p_i(t) - p_i(\tau) - a, t-\tau)]$. As the $p_i(t)$ are continuous, choosing γ sufficiently small guarantees that the derivatives of K_N^T appearing in the definition of θ_i are always evaluated away from their singularity, in particular they are bounded and locally Lipschitz-continous, which implies with the local Lipschitz constant λ

$$|\theta_1(t,\tau) - \theta_2(t,\tau)| \leq \lambda |p_1(t) - p_1(\tau) - p_2(t) + p_2(\tau)| \leq 2\lambda \|p_1(t) - p_2(t)\|_{C([T,T+\gamma])}.$$

Taking the absolute value on both sides of (4.9) and using Lemma 4.6 implies

$$\begin{aligned} |\Lambda_1(t) - \Lambda_2(t)| \leq{} & C_{20} \left(\|f_1^0 - f_2^0\|_{L^2((-L,L))} + \|f_1^0 - f_2^0\|_{H^4(I)}\right) + C_{21}\|p_1(t) - p_2(t)\|_{C([T,T+\gamma])} \\ & + C_{22} \int_T^t |\Lambda_1(\tau) - \Lambda_2(\tau)|\, d\tau, \end{aligned}$$

so that Gronwall's lemma implies, together with (A3) and (A6), yields

$$|\Lambda_1(t) - \Lambda_2(t)| \leq C_{23}(\delta_p + \delta_\Lambda) e^{C_{22} t}. \tag{4.10}$$

Next we exploit the fact that $f_i(p_i(t), t) = 0$ by taking the time derivative, which gives

$$0 = \frac{d}{dt} f_i(p_i(t)) = \dot{p}_i(t) \partial_x f_i(p_i(t), t) + \partial_t f_i(p_i(t), t), \quad i = 1, 2.$$

Subtracting the above equation for $i = 1$ and $i = 2$ respectively, using the definition of Λ_i and integrating in time we obtain, for $T \leq t \leq T + \gamma$

$$\begin{aligned} p_1(t) - p_2(t) = & \int_T^t \left(\frac{\Lambda_2(s)\partial_t f_1(p_1(s), s) - \Lambda_1(s)\partial_t f_2(p_2(s), s)}{\Lambda_1(s)\Lambda_2(s)} \right) ds \\ & + (p_1(T) - p_2(T)) \end{aligned} \tag{4.11}$$

Denoting by $\underline{\Lambda} = \inf_{T \leq s \leq T+\gamma} \Lambda_1(s)\Lambda_2(s)$ and using (A3) this yields

$$\begin{aligned} |p_1(t) - p_2(t)| \leq{} & \frac{1}{\underline{\Lambda}} \int_T^t |(\Lambda_2(s) - \Lambda_1(s))\partial_t f_1(p_1(s), s)|\, ds \\ & + \frac{1}{\underline{\Lambda}} \int_T^t |\Lambda_1(s)(\partial_t f_1(p_1(s), s) - \partial_t f_2(p_1(s), s))|\, ds \\ & + \frac{1}{\underline{\Lambda}} \int_T^t |\Lambda_1(s)(\partial_t f_2(p_1(s), s) - \partial_t f_2(p_2(s), s))|\, ds + \delta_p. \end{aligned}$$

As a consequence of (4.6), $\partial_t f_i(\cdot, t)$ is bounded and Lipschitz continuous. Thus using (4.10), (A3) and once more (4.6) applied to $\partial_t f_1(p_1(t), t) - \partial_t f_2(p_1(t), t)$ (and together with (A6)) finally yields the assertion. $\square$

5 Numerical Simulation

We conclude by illustrating the proposed methodologies and confirming the obtained analytic results with various computational experiments. All simulations are performed on the domain $[-L, L]$, which is split into N intervals of length h. The discrete grid points are denoted by $x_i = ih$. We compute solutions at discrete times $t^k = k\Delta t$, where Δt is the discrete time step.

However we will omit all full time-discrete expressions in the following, to enhance readability.

The reconstruction of the buyer-vendor distribution is based on piecewise linear basis functions. Let V_h denote the space of piecewise linear basis functions ϕ_j, which satisfy $\phi_j(x_i) = \delta_{ij}$. We wish to reconstruct $\hat{f} \in V_h$, which is given by $\hat{f}(x,T) = \sum_{j=1}^{J} \hat{f}_j \phi_j(x)$ using the duality estimates (3.2).

Data generation: We solve the transformed LL model (2.2) for a given initial buyer-vendor distribution f_0. In doing so we transform the initial distribution f_0 via (2.1), and compute the solution to the heat equation (2.2) using an implicit in time discretization. The returned discrete price $p_i = p_i(i\Delta t)$ corresponds to the zero levelset of the buyer-vendor distribution F_i (computed via linear interpolation). Note that we use a finer spatial and temporal discretization to generate the data than in the subsequent reconstruction.

Steepest descent: We solve (3.7) and the corresponding problem on $\Omega_{\triangleright}$ and $\Omega_{\triangleleft}$ using steepest descent. In doing so, we compute the variational derivatives of (3.8) and obtain the first order optimality system (3.9) as well as the updates for the controls u_1 and u_2. The detailed steps are outlined in the While-Loop of Algorithm 1. Here the parameter $\beta > 0$ is the step size of the steepest descent update. Note that we transform the computational domains $\Omega_{\triangleleft}$ and $\Omega_{\triangleright}$ to $[0,1]$ as discussed in Section 2.3 in all simulations. We solve the forward as well as the adjoint equations using an implicit in time discretization and piecewise linear basis functions in space.

Identifiabilty for different inital conditions.

In the first experiment we set $L = 0.5$ and the final time to $T = 0.3$. We split the spatial domain $[-0.5, 0.5]$ into 200 elements and the time interval $[0, 0.3]$ into 100 time steps. The initial datum is set to

$$f_0(x) = (x + 0.75)(x - 0.65)(x - 0.05). \tag{5.1}$$

We approximate the final buyer-vendor distribution using $J = 20$ basis functions. Furthermore we choose the following parameters

$$\alpha = 0.1,\ \beta = 0.25, \text{max. iterations} = 250 \text{ and max error} = 10^{-4}.$$

Figure 2 shows the reconstructed and computed function F (the latter computed by solving the heat equation (2.2) with the transformed initial datum F_0). We observe a good agreement, with small artefacts at the boundary and the buyer-vendor interface. The corresponding controls are shown in Figure 3.

Next we choose a slightly different initial datum, in particular

$$f_0(x) = (x + 0.75)(x - 0.65)(x + 0.05)$$

In this case the price is not monotone, see Figure 4. We observe that the quality of the reconstructions is comparable to the one of the previous example. However the step size of the steepest descent method β had to be decreased to 0.1 to ensure convergence.

Algorithm 1: Reconstruction of $\hat{f}(x,T)$.

Given: Price p_i and transaction rate λ_i at discrete times $t_i = i\Delta t$

for $i = 1 \dots J$ **do**

  if $x_j < p(T)$ *where* $\phi_i(x_j) = \delta_{ij}$ **then**

    $\psi_1(x) = \phi_i(x)$

  else

    $\psi_2(x) = \phi_i(x)$

  k = 0

  while $k <$ *max. iterations and convergence criterion is not satisfied*

    Adjoint equ.: Given $u_1^{i,k}(t)$, $u_2^{i,k}(t)$, $\psi_1(x)$ and $\psi_2(x)$ solve

$$
\begin{aligned}
&\partial_t \Phi_1^{i,k}(x,t) + \partial_{xx}\Phi_1^{i,k}(x,t) = 0 && \partial_t \Phi_2^{i,k}(x,t) + \partial_{xx}\Phi_2^{i,k}(x,t) = 0\\
&\partial_x \Phi_1^{i,k}(-L,t) = 0 && \partial_x \Phi_2^{i,k}(L,t) = 0\\
&\Phi_1^{i,k}(p(t),t) = u_1^{i,k}(t) && \Phi_2^{i,k}(p(t),t) = u_2^{i,k}(t)\\
&\Phi_1^{i,k}(x,T) = \psi_1(x) && \Phi_2^{i,k}(x,T) = \psi_2(x)
\end{aligned}
$$

    Forward equ.: For $G_1^{i,k}(x,0) = -\Phi_1^{i,k}(x,0)$ and $G_2^{i,k}(x,t) = -\Phi_2^{i,k}(x,0)$ solve

$$
\begin{aligned}
&-\partial_t G_1^{i,k}(x,t) + \partial_{xx} G_1^{i,k}(x,t) = 0 && -\partial_t G_2^{i,k}(x,t) + \partial_{xx} G_2^{i,k}(x,t) = 0\\
&\partial_x G_1^{i,k}(-L,t) = 0 && \partial_x G_2^{i,k}(L,t) = 0\\
&G_1^{i,k}(p(t),t) = 0 && G_2^{i,k}(p(t),t) = 0
\end{aligned}
$$

    Update controls $u_1^{i,k} = u_1(t)$ and $u_2^{i,k} = u_2(t)$

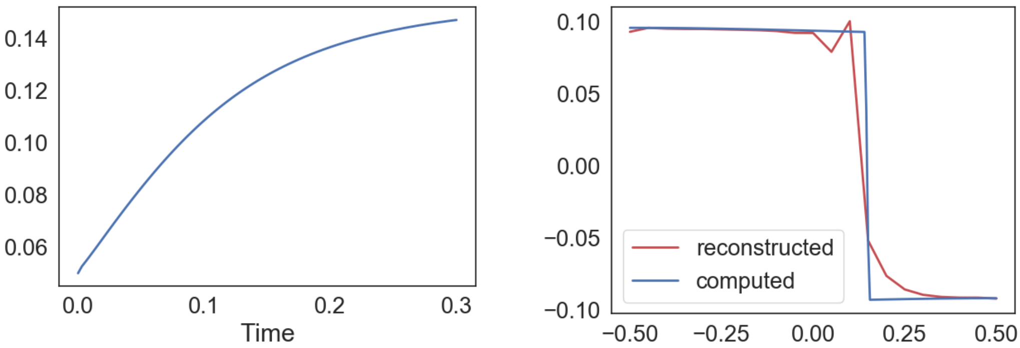
$$
\begin{aligned}
u_1^{i,k+1}(t) &= u_1^{i,k}(t) - \beta\left(\alpha u_1^{i,k}(t) + \partial_x G_1^{i,k}(p(t),t)\right)\\
u_2^{i,k+1}(t) &= u_2^{i,k}(t) - \beta\left(\alpha u_2^{i,k}(t) - \partial_x G_2^{i,k}(p(t),t)\right).
\end{aligned}
$$

    k = k+1

Reconstruct solution $\hat{f}(x) = \sum_j \hat{f}_j \phi_j(x)$:

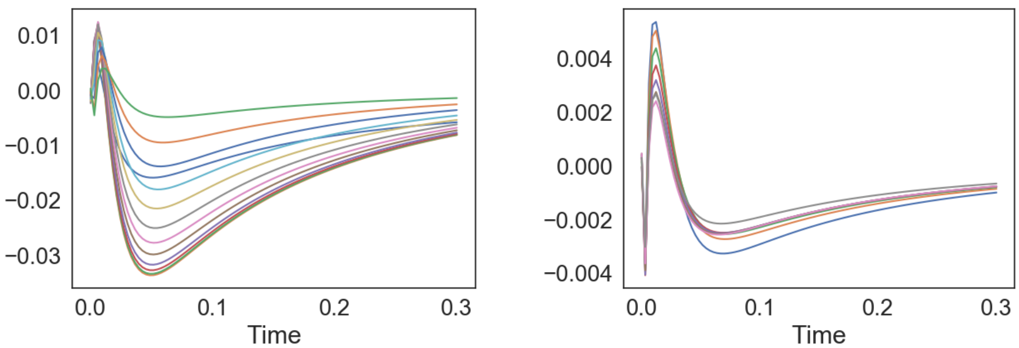
$$
\sum_j \int_{-L}^{p(T)} \hat{f}_j \phi_j(x)\phi_i(x)dx = \int_{-L}^{p(\varepsilon)} f(x,0)\Phi_1^i(x,0)dx + \int_{\varepsilon}^{T} \Lambda(t)(\Phi_1^i(p(t)-a) - u_1^i(t))dt
$$

$$
\sum_j \int_{p(T)}^{L} \hat{f}_j \phi_j(x)\phi_i(x)dx = \int_{p(\varepsilon)}^{L} f(x,0)\Phi_2^i(x,0)dx + \int_{\varepsilon}^{T} \Lambda(t)(u_2^i(t) - \Phi_2^i(p(t)-a))dt
$$



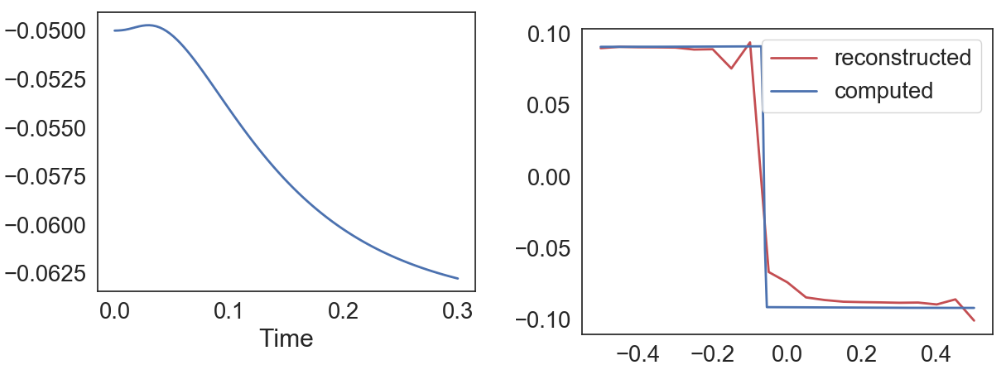
(a) Price dynamics (b) Reconstruction

Figure 2: Left: Evolution of the price $p = p(t)$; Right: Reconstructed and computed buyer-vendor distributions F.



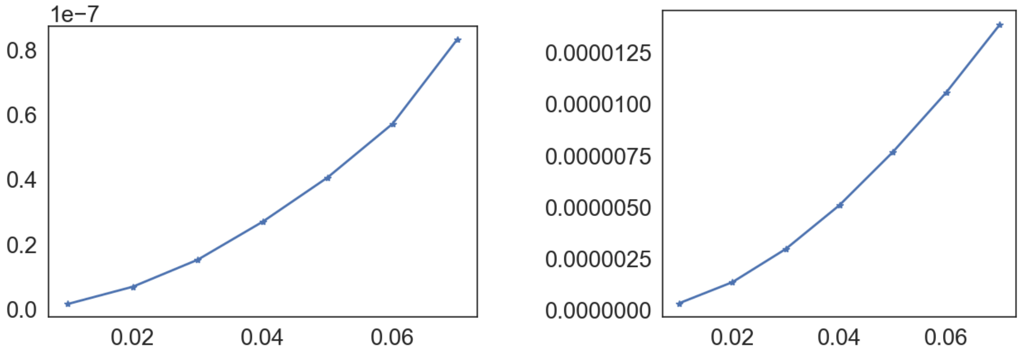
(a) Controls u_1 (b) Controls u_2

Figure 3: Evolution of the controls u_1 and u_2.



(a) Price dynamics (b) Reconstruction

Figure 4: Left: Evolution of the price $p = p(t)$; Right: Reconstructed and computed buyer-vendor distributions F.



(a) Difference in the L^2 norm of the controls (b) Difference in the C^1 norm of the reconstructions

Figure 5: Difference in the controls and the reconstructions for different values of δ.

Stability of $\hat{f}(x,T)$

Next we are interested in the stability of the reconstruction $\hat{f}(x,T)$ with respect to perturbations in the price. Lemma 4.3 and in particular (4.5) state that the difference in the reconstructions is bounded by the difference in the prices and transaction rates. We consider the following perturbation of the unperturbed price $p_{\delta,0}$:

$$p_{\delta,k} = p_{\delta,0} + k\delta \sin(\pi t)$$

for $k = 1,\dots,K$. Note that the pertrubed price is still in C^1 and that $p_{\delta,0}(0) = p_{\delta,k}(0)$ for all $k = 1,\dots K$. We use the first initial datum (5.1) and set $K = 7$. All other parameters are the same as in the first example. Figure 5 illustrates the linear increase of the error in the controls and the reconstruction as the noise level increases.

Predicting price dynamics

We conclude with an example where we use the reconstructed buyer-vendor distribution $f(x,T)$ to estimate future price dynamics and illustrate the influence of noise in those. We consider an initial datum of the form

$$f_0(x) = (x+1.05)(x-1.25)(x-0.1),$$

on the domain $\Omega = [-1,1]$, that is $L = 1$. The computational domain is split into 200 intervals, the time horizon $[0,0.4]$ into 100 time steps. The parameters used in the reconstruction Algorithm 1 are set to

$$\alpha = 0.05,\ \beta = 0.1,\ \text{ max iterations } 150,\ \text{ and max error } 10^{-4}.$$

The final buyer vendor distribution is reconstructed using $J = 50$ basis functions. We use this reconstruction to compute the price dynamics for times $[0.4,0.6]$. Figure 6 illustrates the price dynamics for different perturbations. Each color corresponds to a perturbation of the form

$$p_{\delta,k} = p_{\delta,0} + k\delta \sin\left(\frac{2\pi t}{\sqrt{k\delta}}\right) \tag{5.2}$$

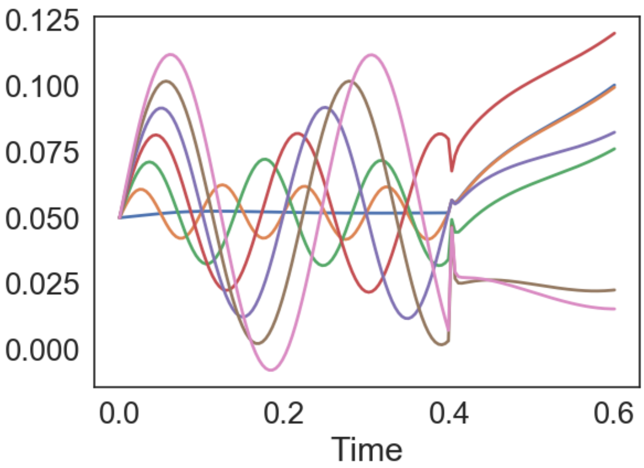


Figure 6: Influence of the perturbations on future price dynamics

for $k = 1, \dots 6$ and a noise level $\delta = 0.01$. Up to time $t = 0.4$ we see sinusoidal price fluctuations due to (5.2). Then the reconstructed profile $f(x, T)$ is used to compute the future price dynamics in the time interval $[0.4, 0.6]$. Here we observe a square root like behavior for all curves towards a stationary price as predicted by the theory. However the stationary price of each curve is different. In [18] Markowich et al. showed that periodic fluctuations in the masses of buyers and vendors lead to periodic fluctuations in the price. The chosen perturbation (5.2) can be related to fluctuations in the buyer vendor masses - these masses determine the stationary price and therefore each curve converges to a different value. Note that the jump at $t = 0.4$ can be explained by the different discretizations used (200 elements in Algorithm 1 vs. 50 elements to compute the future price dynamics).

6 Summary and Outlook

We studied a data assimilation problem for a parabolic nonlinear free boundary problem. This partial differential equation describes the evolution of the price, that is the free boundary, in a large economic market. We developed an analytical and computational framework for the corresponding data assimilation problem, which is based on a previous work by Puel et al., see [20]. The free boundary splits the original problem into two parts, each of them defining a separate optimal control problem. We discussed analytic properties of the respective problems and derived stability estimates for the controls and reconstructed unknown buyer-vendor distribution in the presence of noise. Finally we confirmed and illustrated our results with computational experiments.

We believe that the developed framework provides the basis for more general data assimilation problems in price formation. In [2] Burger et al. considered a Boltzmann type price formation model, which allows for more complex trading mechanisms. This problem is a system of nonlocal reaction-diffusion equations on the whole domain, where multiple prices (even with continuous distribution) and transaction rates can appear. Analogous questions can be asked for this problem if only the expectation of the price is to be predicted, but the problem could also be extended to a stochastic distribution of the price.

Acknowledgements

MB acknowledges support by ERC via Grant EU FP 7 - ERC Consolidator Grant 615216 LifeInverse. MTW acknowledges financial support from the Austrian Academy of Sciences ÖAW via the New Frontiers Grant NST-001 and the EPSRC via the First Grant EP/P01240X/1.

References

[1] A. Alphonse, C. M. Elliott, and B. Stinner. An abstract framework for parabolic pdes on evolving spaces. *Portugaliae Mathematica*, 72:1–46, 2015.

[2] M. Burger, L.A. Caffarelli, P.A. Markowich, and M.-T. Wolfram. On a Boltzmann-type price formation model. 469(2157):20130126, 2013.

[3] M. Burger, L.A. Caffarelli, P.A. Markowich, and M.-T. Wolfram. On the asymptotic behavior of a Boltzmann-type price formation model. *Communications in Mathematical Sciences*, 12(7), 2014.

[4] L.A. Caffarelli, P.A. Markowich, and J.-F. Pietschmann. On a price formation free boundary model by Lasry and Lions. *Comptes Rendus Mathematique*, 349(11-12):621 – 624, 2011.

[5] L.A. Caffarelli, P.A. Markowich, and M.-T. Wolfram. On a price formation free boundary model by Lasry and Lions: the Neumann problem. *Comptes Rendus Mathematique*, 349(15):841–844, 2011.

[6] J.R. Cannon. *The one-dimensional heat equation*, volume 23 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1984. With a foreword by Felix E. Browder.

[7] D. Chae, O.Y.. Imanuvilov, and S.M. Kim. Exact controllability for semilinear parabolic equations with neumann boundary conditions. *Journal of Dynamical and Control Systems*, 2(4):449–483, 1996.

[8] L. Chayes, M. del Mar González, M.P. Gualdani, and I. Kim. Global existence and uniqueness of solutions to a model of price formation. *SIAM Journal on Mathematical Analysis*, 41(5):2107–2135, 2009.

[9] C. Clason and P. Hepperger. A forward approach to numerical data assimilation. *SIAM Journal on Scientific Computing*, 31(4):3090–3115, 2009.

[10] R. Cont and M. S. Mueller. A stochastic PDE model for limit order book dynamics. *arXiv e-prints*, page arXiv:1904.03058, Apr 2019.

[11] M. del Mar Gonzàlez and M. P. Gualdani. Asymptotics for a free boundary model in price formation. *Nonlinear Analysis: Theory, Methods & Applications*, 74(10):3269 – 3294, 2011.

[12] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[13] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, 2008.

[14] M. V. Klibanov. Inverse problems and carleman estimates. *Inverse problems*, 8(4):575, 1992.

[15] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, Providence, rhode island, 1968.

[16] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[17] J. L. Lions. 1. pointwise control for distributed systems. In H.T. Banks, editor, *Control and Estimation in Distributed Parameter Systems*, chapter 1, pages 1–39. SIAM, 1992.

[18] P. Markowich, J. Teichmann, and M. Wolfram. Parabolic free boundary price formation models under market size fluctuations. *Multiscale Modeling & Simulation*, 14(4):1211–1237, 2016.

[19] P.A. Markowich, N. Matevosyan, J.-F. Pietschmann, and M.-T. Wolfram. On a parabolic free boundary equation modeling price formation. *Mathematical Models and Methods in Applied Sciences*, 19(10):1929–1957, 2009.

[20] J.-P. Puel. Une approche non classique d'un problme d'assimilation de donnes. *Comptes Rendus Mathematique*, 335(2):161 – 166, 2002.

[21] J.-P. Puel. A nonstandard approach to a data assimilation problem and tychonov regularization revisited. *SIAM Journal on Control and Optimization*, 48(2):1089–1111, 2009.

[22] F. Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications*, volume 112. American Mathematical Soc., 2010.

[23] M. Yamamoto. Carleman estimates for parabolic equations and applications. *Inverse problems*, 25(12):123013, 2009.