HAL
open science

BUILDING ARBITRAGE-FREE IMPLIED VOLATILITY: SINKHORN'S ALGORITHM AND VARIANTS

Hadrien de March, Pierre Henry-Labordere

To cite this version:

Hadrien de March, Pierre Henry-Labordere. BUILDING ARBITRAGE-FREE IMPLIED VOLATILITY: SINKHORN'S ALGORITHM AND VARIANTS. 2019. hal-02011533

HAL Id: hal-02011533

https://hal.science/hal-02011533v1

Preprint submitted on 8 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BUILDING ARBITRAGE-FREE IMPLIED VOLATILITY: SINKHORN'S ALGORITHM AND VARIANTS

HADRIEN DE MARCH AND PIERRE HENRY-LABORDÈRE

Abstract. We consider the classical problem of building an arbitrage-free implied volatility surface from bid-ask quotes. We design a fast numerical procedure, for which we prove the convergence, based on the Sinkhorn algorithm that has been recently used to solve efficiently (martingale) optimal transport problems.

1. Introduction

Building arbitrage-free implied volatility surfaces from bid-ask quotes is a long-standing issue. In particular, this is needed for market-makers in equity Vanillas. This is also needed for pricing exotic options when using risk-neutral models calibrated to Vanillas, as for the local volatility model [9] or for local stochastic volatility models [19]. In this purpose, various approaches have been considered. We review in the next section some of them and highlight their main drawbacks. The definite answer should be able to:

(1) produce calendar/butterfly arbitrage-free surfaces.
(2) fit market quotes perfectly within bid/ask spreads.
(3) fit smiles before earnings (with Mexican hat-shape curves).
(4) fit quickly.

1.1. **Review of literature.** For completeness, we recall that the market price of a call option $\mathcal{C}(T,K) \in [S_0,\infty)$ with maturity T and strike K is quoted in terms of an implied volatility $\sigma_{\mathrm{BS}}(T,K)$ defined as the constant volatility $\sigma_{\mathrm{BS}}(T,K) := \sigma$ such that $\mathcal{C}(T,K) = \mathrm{BS}(S_0,K,\sigma\sqrt{T})$ where S_0 is the spot price value at $t=0$ and BS denotes the Black-Scholes formula:

$$\mathrm{BS}(S_0,K,\omega) := S_0 N(d_+) - K N(d_-).$$

Here $d_\pm = \frac{\ln \frac{S_0}{K}}{\omega} \pm \frac{\omega}{2}$ and $N(x)$ is the standard normal cumulative distribution function. As $\mathrm{BS} \in [S_0,\infty)$ is strictly increasing in ω, the implied volatility is unique. In the following, for ease of notations, we assume zero rates/dividends (see however remark (2.3) for explanations how to include exactly cash/yield dividends (and deterministic rates) in this framework).

1.1.1. *SVM-based parameterization.* We consider the implied volatility associated to a stochastic volatility model (in short SVM), depending on some parameters: initial volatility, spot-volatility

correlation, volatility-of-volatility, etc.... For example, one can consider an SVM, defined by an homogeneous Itô diffusion:

$$dS_t = C(S_t)a_t dW_t, \quad da_t \quad = \quad (\cdots)dt + \sigma(a_t)dZ_t, \quad d\langle Z, W\rangle_t = \rho dt.$$

As coming from a risk-neural model (i.e., $S_\cdot$ is a (local) martingale – see [14] for sufficient and necessary conditions on the coefficients of the diffusion with $C(s) := s$ for imposing that S is not only a local martingale but a true martingale), the resulting implied volatility $,\sigma_{\rm BS}(\cdot,\cdot)$, for which $\mathbb{E}[(S_T - K)^+] = {\rm BS}(S_0, K, \sigma_{\rm BS}(T,K)\sqrt{T})$, is arbitrage-free. In practice, the implied volatility can not be derived in closed-form and therefore the calibration of the parameters of the SVM on market prices can be quite time-consuming. In order to speed up this optimization, one can rely on the approximation of the implied volatility in the short-maturity regime. At the first-order in the maturity T, one can derive a generic formula [14], obtained by using short-time asymptotics of the heat kernel on Cartan-Hadamard manifolds, for which the cut-locus is empty:

$$\begin{aligned} \sigma_{\rm BS}(T,K) \quad &\underset{T\to 0}{\sim} \quad \frac{\ln \frac{K}{S_0}}{\int_{S_0}^K \frac{dx}{a^*(x)}} \left(1 + a_1(K)T + O(T^2)\right), \\ a^*(x) \quad &:= \quad {\rm argmin}_a d_{\rm geo}(x, a|S_0, a_0), \end{aligned} \tag{1.1}$$

where the geodesic distance $d_{\rm geo}$ is

$$d_{\rm geo}(y_2, x_2|y_1, x_1) := \int_{y_1}^{y_2} \frac{F(y')}{\sqrt{F(y') - C^2}} dy',$$

with C defined by the equation $x_2 - x_1 = \int_{y_1}^{y_2} \frac{C}{\sqrt{F(y')-C^2}} dy'$, and $F(y) := \frac{2}{a(y)^2(1-\rho^2)}$, with the new coordinates $x := \int_{S_0}^S \frac{dz}{C(z)} - \rho \int_{a_0}^a \frac{u}{\sigma(u)} du$ and $y := \sqrt{1-\rho^2} \int_{a_0}^a \frac{u}{\sigma(u)} du$. The lengthly expression for $a_1(K)$ is not reported and can be found in [14]. As an example, one can cite the SABR parameterization for which $C(S) := S^\beta$ with $\beta \in [0,1)$ and a_t is a log-normal process. The resulting manifold is the $2d$ hyperbolic space $\mathbb{H}^2$. Let us remark that similar formulas can be also derived using large deviations (see [10] for extensive references).

By construction, the implied volatility is arbitrage-free in strike as the parametrization comes from a risk-neutral model. However, the maturity T should be "small" in order to preserve the validity of our approximate formula (1.1). The arbitrability in maturity is not ensured as the calibration is performed by considering separately each time slice. Moreover, as our formula depends on a finite number of parameters, it is not possible to match exactly market prices. From a numerical point of view, the calibration involves a non-convex optimization, which is not guaranteed to converge. This solution only solves (4) and partially (1).

1.1.2. *Parametric form.* Another approach is to start directly with a parametrization of the implied volatility. As an example, commonly used by practitioners, we have the SVI parametrization [12]

$$\sigma_{BS}(T,K) = a + b\left(\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}\right),$$

depending on five parameters a, b, ρ, m and σ. Note that this parametrization can be linked with the large maturity limit of the implied volatility in the Heston model. Despite its simplicity, the arbitrage-freeness in strike and maturity is not guaranteed, see however [11] for some conditions

on the term-structures of the parameters (in maturity) which ensure an arbitrage-free surface [12]. These limitations restrict the space of admissible parameters and therefore this solution only solves (4) and partially (1).

1.1.3. *Discrete local volatility.* One approach to impose the arbitrage-freeness in strike and maturity is to start (again) with a non-homogenous risk-neutral model. One can use a discrete local volatility [1]. Given a time grid of expiries $0 := t_0 < t_1 < \cdots < t_n$, call prices $c(t_i, \cdot)$ at time t_{i+1} are then taken to be solutions of the ODE:

$$\left[1 - \frac{1}{2}\Delta t_i \sigma_i(k)^2 \partial_k^2\right] c(t_{i+1}, k) = c(t_i, k), \quad c(0, k) = (S_0 - k)^+.$$

By using for $\sigma_i(\cdot)$ a piecewise constant function, we can try to match market prices of call options. As pointed in [18], "this method uses a fully implicit finite-difference scheme to compute the probability density of the underlying, stepping forward in time and calibrating model parameters by a least-squares algorithm. Since the size of time step is determined by market quotes, it cannot be reduced arbitrarily, so that, while very instructive, this method clearly has limited accuracy". For example, with this algorithm, we were not able to calibrate equity Vanillas exhibiting a Mexican hat form (see Figure 1), just before earning dates. Some improvements have been considered in [18].

1.2. **Contents.** In this paper, we will build a solution satisfying (1-2-3-4) by construction. The conditions (1-2-3) are automatically (and exactly) satisfied as we construct a non-parametric density fitting the Vanillas. Our approach is close in spirit to the "Weighted Monte-Carlo approach" based on an entropic penalisation as introduced in [2]. However, our approach takes in account the calendar spread requirement and therefore is able to produce (arbitrage-free) Vanillas at different maturities increasing in the convex order. Furthermore, by relying on the Sinkhorn's algorithm that has been popularized recently for solving quickly optimal transportation problems ([7], [20]), we present a Sinkhorn's algorithm compatible with the convex order property (see also [13], [8]). The convergence of our algorithm is then proved (see Theorem 4.5) with a fast decay rate and therefore our numerical scheme solves (4). We conclude with numerous examples of fitting to Equity Vanillas for various stocks and indices.

2. Axiomatics: Formulation

Prices of call options for different maturities $t_1 \leq \cdots \leq t_n$ and different strikes are quoted on the market. We denote by $\mathcal{C}_i^K$ the market prices of maturity t_i and strike $K \in \mathcal{K}_i$. The set $\mathcal{K}_i$ corresponds to the strikes $K_i^1 < \cdots < K_i^{n_i}$. Building an arbitrage-free implied volatility is equivalent to find a martingale probability measure $\mathbb{P}^*$ in $\mathbb{R}_+^n$ that matches (exactly) this market prices: $\mathbb{P}^*$ should belong to the convex set

$$\mathcal{M}_n = \left\{\mathbb{P} \,:\, \mathbb{E}^{\mathbb{P}}[(S_{t_i} - K)^+] = \mathcal{C}_i^K, \quad \forall K \in \mathcal{K}_i,\ \mathbb{E}^{\mathbb{P}}[S_{t_i} | S_0, \cdots, S_{t_{i-1}}] = S_{t_{i-1}}, \quad i = 1, \cdots, n\right\}.$$

For use below, we set $\mathcal{C}_i^j := \mathcal{C}_i^{K_i^j}$ and

$$\begin{aligned}
\mathrm{VS}_i^j &:= \frac{\mathcal{C}_i^{j-1} - \mathcal{C}_i^j}{K_j^i - K_i^{j-1}} \quad 1 \le j \le n_i, \\
\mathrm{VS}_i^0 &:= 1, \\
\mathrm{CVS}_{i_1,i_2}^{j_1,j_2} &:= C_{i_2}^{j_2} - C_{i_1}^{j_1}, \\
\mathrm{CBS}_{i,i_1,i_2}^{j,j_1,j_2} &:= \frac{CVS_{i_1,i}^{j_1,j}}{K_{i_1}^{j_1} - K_i^j} - \frac{CVS_{i,i_2}^{j,j_2}}{K_i^j - K_{i_2}^{j_2}}.
\end{aligned}$$

For completeness, we cite the following result that gives necessary and sufficient conditions for arbitrage-freeness:

Lemma 2.1 (see [5, 6] for proofs)**.** *$\mathcal{M}_n$ is non-empty if and only if for all $i = 1, \cdots, n$*

(1)

$$\begin{aligned}
&\mathcal{C}_i^j \ge 0, \quad 0 \le j \le n_i, \\
&\mathrm{VS}_i^j \in [0,1], \quad 1 \le j \le n_i, \\
&\mathrm{VS}_i^j > 0 \quad \text{if} \quad \forall\, 1 \le j \le n_i, \quad \text{and } C_i^{j-1} > 0.
\end{aligned}$$

(2) *$\forall i_1, i_2 \in [1,n]$ s.t. $i_1 < j_2$, $\forall j_1 \in [0, n_{i_1}]$, $\forall j_2 \in [0, m_{i_2}]$*

$$\begin{aligned}
&\mathrm{CVS}_{i_1,i_2}^{j_1,j_2} \ge 0, \quad \text{if} \quad K_{i_1}^{j_1} \ge K_{i_2}^{j_2}, \\
&\mathrm{CVS}_{i_1,i_2}^{j_1,j_2} > 0, \quad \text{if } K_{i_1}^{j_1} > K_{i_2}^{j_2} \text{ and } \mathcal{C}_{i_1}^{j_1} > 0.
\end{aligned}$$

(3) *$\forall i, i_1, i_2 \in [1,n]$ s.t. $i \le i_1$ and $i \le i_2$, $\forall j \in [0, n_i]$, $\forall j_2 \in [0, n_{i_2}]$ s.t. $K_{i_1}^{j_1} < K_i^j < K_{i_2}^{j_2}$:*

$$\mathrm{CBS}_{i,i_1,i_2}^{j,j_1,j_2} \ge 0.$$

Markovian solutions. As a simplification, we could assume that $\mathbb{P}^*$ should satisfy a Markov property and therefore belongs instead to the subset of $\mathcal{M}_n$:

$$\mathcal{M}_n^{\mathrm{Markov}} = \{\mathbb{P} \quad : \quad \mathbb{E}^{\mathbb{P}}[(S_{t_i} - K)^+] = \mathcal{C}_i^K, \quad \forall K \in \mathcal{K}_i, \quad \mathbb{E}^{\mathbb{P}}[S_{t_i} | S_{t_{i-1}}] = S_{t_{i-1}}, \quad i = 1, \cdots, n\}.$$

Lemma 2.2. *$\mathcal{M}_n^{\mathrm{Markov}}$ is non-empty if and only $\mathcal{M}_n$ is non-empty. In particular if the market data $(\mathcal{C}_i)_{1 \le i \le n}$ are arbitrage-free, there can be attained by a martingale measure in $\mathcal{M}_n^{\mathrm{Markov}}$.*

Proof.

$\Longrightarrow$: obvious.

$\Longleftarrow$ Take $\mathbb{P} \in \mathcal{M}_n$. Then by disintegration, define the marginals $\mathbb{P}^{i-1}$ and $\mathbb{P}^i$, which are in the convex order. From Kellerer's theorem, we can build a martingale measure $\mathbb{P}^{i-1,i}$ with marginals $\mathbb{P}^{i-1}$ and $\mathbb{P}^i$ (see e.g. [17] for an explicit construction). By gluing these measures, we get an element in $\mathcal{M}_n^{\mathrm{Markov}}$.

2.1. **Sequential construction.** From the Markov property, an element $\mathbb{P} \in \mathcal{M}_n^{\text{Markov}}$ could be written as

$$\mathbb{P}(ds_1, \cdots, ds_n) = \mathbb{P}^1(ds_1) \prod_{i=2}^{n} \mathbb{P}^{i-1,i}(ds_i | ds_{i-1}),$$

where the probability $\mathbb{P}^1$ and $(\mathbb{P}^{i-1,i})_{i=1,\cdots,n}$ are constructed as follows:

(1) We choose an $\mathbb{P}^1 \in \mathcal{M}_1^{\text{Markov}}$ (a specific example is constructed in Section 3) with

$$\mathcal{M}_1^{\text{Markov}} = \{\mathbb{P} \quad : \quad \mathbb{E}^{\mathbb{P}}[(S_{t_1} - K)^+] = \mathcal{C}_1^K, \quad \forall K \in \mathcal{K}_1, \quad \mathbb{E}^{\mathbb{P}}[S_{t_1}|S_0] = S_0\}.$$

(2) We choose an $\mathbb{P}^{1,2} \in \mathcal{M}_{1,2}^{\text{Markov}}$ (a specific example is constructed in Section 4) with

$$\mathcal{M}_{1,2}^{\text{Markov}}(\mathbb{P}^1) = \{\mathbb{P} \ : \ S_{t_1} \overset{\mathbb{P}}{\sim} \mathbb{P}^1, \quad \mathbb{E}^{\mathbb{P}}[(S_{t_2} - K)^+] = \mathcal{C}_2^K, \quad \mathbb{E}^{\mathbb{P}}[S_{t_2}|S_{t_1}] = S_{t_1}\}.$$

From $\mathbb{P}^{1,2} \in \mathcal{M}_{1,2}^{\text{Markov}}$, we define $\mathbb{P}^2$ as

$$\mathbb{P}^2(ds_2) = \int \mathbb{P}^{1,2}(ds_1, ds_2).$$

(3) We iterate step **(2)** to obtain $(\mathbb{P}^{i-1,i})_{i=3,\cdots,n}$.

2.2. **Adding bid-ask prices.** In practice, market prices are quoted with bid-ask prices. Our discussion can be generalized to this case by replacing $\mathcal{M}_1^{\text{Markov}}$ and $\mathcal{M}_{1,2}^{\text{Markov}}$ by

$$\begin{aligned}
\widetilde{\mathcal{M}}_1^{\text{Markov}} &= \{\mathbb{P} \ : \ \mathcal{C}_1^{K,\text{mid}} \leq \mathbb{E}^{\mathbb{P}}[(S_{t_1} - K)^+] \leq \mathcal{C}_1^{K,\text{ask}}, \quad \forall K \in \mathcal{K}_1, \quad \mathbb{E}^{\mathbb{P}}[S_{t_1}|S_0] = S_0\}. \\
\widetilde{\mathcal{M}}_{1,2}^{\text{Markov}}(\mathbb{P}^1) &= \{\mathbb{P} \ : \ S^1 \overset{\mathbb{P}}{\sim} \mathbb{P}^1, \ \mathbb{E}^{\mathbb{P}}[S_{t_2}|S_{t_1}] = S_{t_1}, \ \mathcal{C}_2^{K,\text{bid}} \leq \mathbb{E}^{\mathbb{P}}[(S_{t_2} - K)^+] \leq \mathcal{C}_2^{K,\text{ask}}, \ \forall K \in \mathcal{K}_2\}.
\end{aligned}$$

We consider this setup in the next sections. The arbitrage-free conditions, which ensure that $\widetilde{\mathcal{M}}_{1,2}^{\text{Markov}}(\mathbb{P}^1)$ is non-empty, are given in [6].

Remark 2.3 (Cash/yield dividends). We assume here that the spot process S_t jumps down by the dividend amounts $D_i(S_{t_i^-}) = \beta_i S_{t_i^-} + \alpha_i$, paid at the dates $0 < t_n < \ldots t_2 < t_1 < T$, and that between dividend dates it follows a diffusion. By setting $S_t = A(t) + B(t)X_t$ (see [15] for formulas for A and B as functions of (α_i, β_i)), one obtains that X_t is a martingale. Call options on S can therefore be written as call options on X. One can then applies our construction to X and deduce then call options on S. Using this mapping, we will assume no dividends/zero rates in the following.

3. Building an element in $\widetilde{\mathcal{M}}_1^{\text{Markov}}$

For the sake of simplicity for the rest of this paper, we denote $S_1 := S_{t_1}$ and $S_2 := S_{t_2}$. An element $\mathbb{P}^* \in \widetilde{\mathcal{M}}_1^{\text{Markov}}$ can be obtained by minimizing a convex lower semi-continuous functional $\mathcal{F}_1$:

$$\inf_{\mathbb{P} \in \widetilde{\mathcal{M}}_1^{\text{Markov}}} \mathcal{F}_1(\mathbb{P}) = \mathcal{F}_1(\mathbb{P}^*), \quad \mathbb{P}^* \in \widetilde{\mathcal{M}}_1^{\text{Markov}}. \tag{3.2}$$

As $\widetilde{\mathcal{M}}_1^{\text{Markov}}$ is weakly compact from Prokhorov's theorem, we deduce that the infimum is attained by an unique $\mathbb{P}^*$.

Remark 3.1 (Reconstructing $\widetilde{\mathcal{M}}_1^{\text{Markov}}$). Note that as $\widetilde{\mathcal{M}}_1^{\text{Markov}}$ is weakly compact, it can be characterized by its extremal points from Krein-Milman's theorem:

$$\widetilde{\mathcal{M}}_1^{\text{Markov}} = \overline{\text{Ext}(\widetilde{\mathcal{M}}_1^{\text{Markov}})}.$$

An extremal point can then be built by maximizing $\mathcal{F}_1$:

$$\sup_{\mathbb{P}\in\widetilde{\mathcal{M}}_1^{\text{Markov}}} \mathcal{F}_1(\mathbb{P}) = \mathcal{F}_1(\mathbb{P}^*), \quad \mathbb{P}^* \in \text{Ext}(\widetilde{\mathcal{M}}_1^{\text{Markov}}).$$

Therefore, by choosing the appropriate functional, one can enumerate in principle all extremal points and therefore reconstruct the space $\widetilde{\mathcal{M}}_1^{\text{Markov}}$ by some linear convex mixtures of the extremal points.

3.1. **Choice of $\mathcal{F}_1$.** We choose $(\omega_K)_{K\in\mathcal{K}_1} \in \mathbb{R}_+^{n_1}$ and consider the regularized Kullback-Leibler functional:

$$\mathcal{F}_1(\mathbb{P}) := \mathbb{E}^{\mathbb{P}}\left[\ln\frac{d\mathbb{P}}{dm_0} - 1\right] + \sum_{K\in\mathcal{K}_1}\frac{1}{2\omega_K}\left(\mathbb{E}^{\mathbb{P}}[(S_1-K)^+] - \mathcal{C}_1^{K,\text{mid}}\right)^2,$$

depending on a prior measure m_0 on $\mathbb{R}^+$, left unspecified for the moment. Let us notice that by introducing dual variables $v_K \in \mathbb{R}$, for each $K \in \mathcal{K}_1$, therefore $\mathcal{F}_1$ may also be written as

$$\mathcal{F}_1(\mathbb{P}) := \mathbb{E}^{\mathbb{P}}\left[\ln\frac{d\mathbb{P}}{dm_0} - 1\right] - \inf_{v\in\mathbb{R}^{\mathcal{K}_1}}\sum_{K\in\mathcal{K}_1} v_K\left(\mathcal{C}_1^{K,\text{mid}} - \mathbb{E}^{\mathbb{P}}[(S_1-K)^+]\right) + \frac{1}{2}v_K^2\omega_K.$$

Proposition 3.2. *The minimization (3.2) is attained by $\mathbb{P}^* \in \widetilde{\mathcal{M}}_1^{\text{Markov}}$ with*

$$\mathbb{P}^*(ds) = m_0(ds)e^{-\sum_{K\in\mathcal{K}_1} V_K^*(s-K)^+ - u_0^* - h_0^*(s-S_0)},$$

where $(V_K^*)_{K\in\mathcal{K}_1}$, u_0, and h_0^* solves the strictly convex unconstrained minimization:

$$\inf_{V\in\mathbb{R}^{\mathcal{K}_1}, u_0, h_0\in\mathbb{R}} \mathcal{G}_1(u_0, h_0, V),$$

$$\begin{aligned} \textit{where} \quad \mathcal{G}_1(u_0,h_0,V) \quad &:= \quad u_0S_0 + \sum_{K\in\mathcal{K}_1} f_1^{K,\text{bid/ask}}(V_K,\omega_K) + \sum_{K\in\mathcal{K}_1} V_K\mathcal{C}_1^{K,\text{mid}} \\ &\quad + \mathbb{E}^{m_0}[e^{-\sum_{K\in\mathcal{K}_1} V_K(S_1-K)^+ - u_0 - h_0(S_1-S_0)}], \end{aligned}$$

$$\begin{aligned} \textit{and} \quad f_i^{K,\text{bid/ask}}(V,\omega) &:= \quad \tfrac{V^2\omega}{2}, && \textit{if } \Delta\mathcal{C}_i^{K,\text{bid}} \le V\omega \le \Delta\mathcal{C}_i^{K,\text{ask}} \\ &:= \quad \Delta\mathcal{C}_i^{K,\text{ask}}V - \tfrac{(\Delta\mathcal{C}_i^{K,\text{ask}})^2}{2\omega}, && \textit{if } \Delta\mathcal{C}_i^{K,\text{ask}} < V\omega \\ &:= \quad \Delta\mathcal{C}_i^{K,\text{bid}}V - \tfrac{(\Delta\mathcal{C}_i^{K,\text{bid}})^2}{2\omega}, && \textit{if } \Delta\mathcal{C}_i^{K,\text{bid}} > V\omega. \end{aligned}$$

Here $\Delta\mathcal{C}_i^{\text{bid/ask}} := \mathcal{C}_i^{\text{bid/ask}} - \mathcal{C}_i^{\text{mid}}$.

This proposition without bid/ask prices originates from [2].

Proof. By introducing dual variables $u^{\text{bid}}, u^{\text{ask}} \in (\mathbb{R}^+)^{\mathcal{K}_1}$ for the inequalities for the call prices at bid and at ask, $\inf \mathcal{G}_1$ may be written as

$$\begin{aligned}\inf \mathcal{G}_1 &= -\inf_{u^{\text{bid}},u^{\text{ask}}\in(\mathbb{R}^+)^{\mathcal{K}_1},v_K\in\mathbb{R},u_0,h_0\in\mathbb{R}} u_0S_0 + \sum_{K\in\mathcal{K}_1} u_K^{\text{ask}}\mathcal{C}_K^{\text{ask}} - u_K^{\text{bid}}\mathcal{C}_K^{\text{bid}} + v_K\mathcal{C}_K^{\text{mid}} \\ &\quad + \frac{1}{2}v_K^2\omega_K + \mathbb{E}^{m_0}[e^{-\sum_{K\in\mathcal{K}}(u_K^{\text{ask}}-u_K^{\text{bid}}+v_K)(S_1-K)^+-u_0-h_0(S_1-S_0)}].\end{aligned}$$

By setting $v := V - u^{\text{ask}} + u^{\text{bid}}$, the function, to be minimized, is equivalent to

$$\begin{aligned} & u_0S_0 + \sum_{K\in\mathcal{K}_1} u_K^{\text{ask}}(\mathcal{C}_1^{K,\text{ask}} - \mathcal{C}_1^{K,\text{mid}}) - u_K^{\text{bid}}(\mathcal{C}_1^{K,\text{bid}} - \mathcal{C}_1^{K,\text{mid}}) + V_K\mathcal{C}_1^{K,\text{mid}} \\ + \quad & \frac{1}{2}(V - u^{\text{ask}} + u^{\text{bid}})^2\cdot\omega + \mathbb{E}^{m_0}\left[e^{-\sum_{K\in\mathcal{K}} V_K(S_1-K)^+-h_0S_1}\right].\end{aligned}$$

We observe that the minimization over u^{ask} and u^{bid} can be exactly performed and we obtain finally an unconstrained optimization over V.

Dependence on the prior. We consider two prior densities $\mathbb{P}_0$ and $\mathbb{P}_0'$. By definition, the vanillas constructed using the two priors satisfy the equations for all $K \in \mathcal{K}_1$:

$$\begin{aligned}\mathcal{C}_1^{K,\text{mid}} + \partial_V f(V_K,\omega_K) - \mathcal{C}^{\text{model}}(K,\mathbb{P}_0) = 0 \\ \mathcal{C}_1^{K,\text{mid}} + \partial_V f(V_K',\omega_K) - \mathcal{C}^{\text{model}}(K,\mathbb{P}_0') = 0\end{aligned}$$

By taking the difference, we get

$$\begin{aligned}|\mathcal{C}^{\text{model}}(K,\mathbb{P}_0') - \mathcal{C}^{\text{model}}(K,\mathbb{P}_0)| &= |\partial_V f(V_K',\omega_K) - \partial_V f(V_K,\omega_K)| \\ &\leq \omega_K|V_K' - V_K|.\end{aligned}$$

3.2. **Numerical examples.** In practice, we take $\omega_K = \Lambda|\mathcal{C}_1^{K,\text{ask}} - \mathcal{C}_1^{K,\text{bid}}|$ with $\Lambda = 0.1$ in our numerical examples. The minimization over V and ω is performed using a modified Newton method and a user-supplied Hessian. In order to have easier computations thanks to the closed formulas displayed in Remark 4.2, we use as a reference measure $m_0(ds_1) := \mathbb{P}_0(ds_1)\mathbf{1}_{s_1\geq 0}$, where $\mathbb{P}_0$ is the Gaussian measure $\mathcal{N}(S_0,\sigma_0^2t_1)$, properly normalized on $\mathbb{R}_+$, and where σ_0 is chosen to minimize the criterion:

$$\inf_{\sigma_0}\sum_{K\in\mathcal{K}_1}\left(\mathbb{E}^{\mathbb{P}_0}[(S_1-K)^+] - \mathcal{C}_1^{K,\text{mid}}\right)^2 = \inf_{\sigma_0}\sum_{K\in\mathcal{K}_1}\left(\text{B}(S_0,t_1,K,\sigma_0) - \mathcal{C}_1^{K,\text{mid}}\right)^2,$$

with

$$\text{B}(s,t,K,\sigma) := \frac{1}{2}(s-K)\text{erf}\left(\frac{K-s}{\sqrt{2}\sigma\sqrt{t}}\right) + \frac{\sigma\sqrt{t}e^{-\frac{(K-s)^2}{2\sigma^2t}}}{\sqrt{2\pi}}.$$

In Figure 1, we show examples of calibration with two stocks (Google & Amazon) near earnings. By construction, the fit is perfect (within the bid/ask spread) and arbitrage-free. In Figure 3.2, we consider two indices (Dax & Euro Stoxx 50).

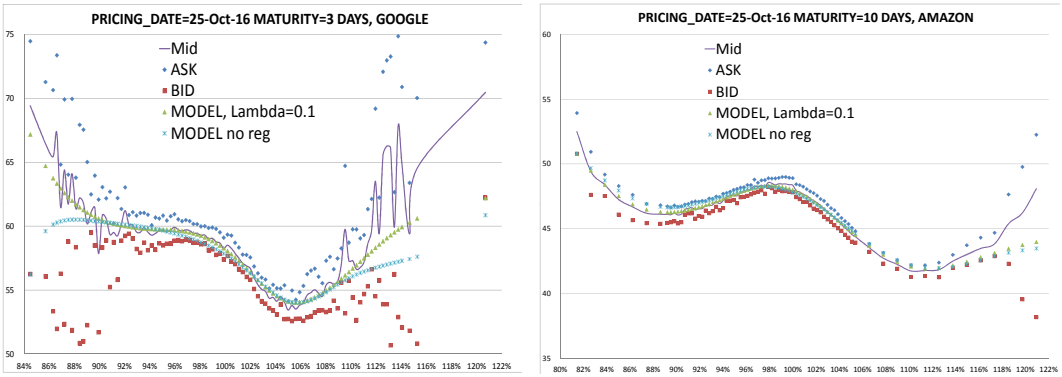

FIGURE 1. Computational time = 0.1 s. Left: GOOGLE. Right: AMAZON. The plots denoted "Model no reg" mean that we have chosen $\Lambda = \infty$.

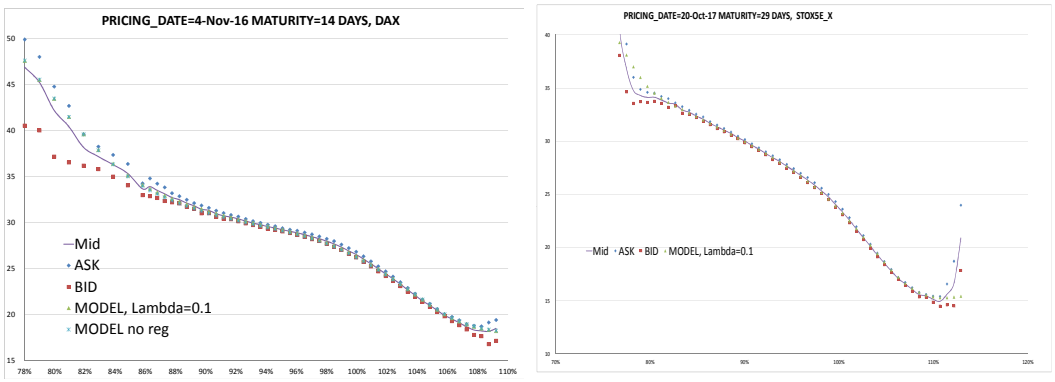

FIGURE 2. Computational time = 0.1 s. Left: DAX. Right: EURO STOXX 50.

4. BUILDING AN ELEMENT IN $\widetilde{\mathcal{M}}_{12}^{\text{Markov}}(\mathbb{P}^1)$

Proceeding as in the previous section, we consider the minimization problem:

$$\inf_{\mathbb{P}\in\widetilde{\mathcal{M}}_{12}^{\text{Markov}}(\mathbb{P}^1)} \mathbb{E}^{\mathbb{P}}\left[\ln\frac{d\mathbb{P}}{dm_0}-1\right]+\sum_{K\in\mathcal{K}_2}\frac{1}{2\omega_K}\left(\mathbb{E}^{\mathbb{P}}[(S_2-K)^+]-\mathcal{C}_2^{K,\text{mid}}\right)^2, \tag{4.3}$$

depending on a prior measure m_0 on $(\mathbb{R}^+)^2$, left unspecified for the moment. Proceeding similarly as in Proposition 3.2 (therefore the proof is not reported, see also [16] for details), we obtain

Proposition 4.1 (see [16] for a proof)**.** *The minimization (4.3) is attained by* $\mathbb{P}^*\in\widetilde{\mathcal{M}}_{12}^{\text{Markov}}(\mathbb{P}^1)$ *with*

$$\mathbb{P}^*(ds_1,ds_2)=m_0(ds_1,ds_2)e^{-\sum_{K\in\mathcal{K}_2}V_K^*(s_2-K)^+-u_1^*(s_1)-h_1^*(s_1)(s_2-s_1)},$$

where u_1^**,* h_1^**, and* V^* *solve the strictly convex unconstrained minimization:*

$$\inf_{V\in\mathbb{R}^{\mathcal{K}_2},u_1\in\mathrm{L}^1(\mathbb{P}^1),h_1\in C_0(\mathbb{R}_+)} \mathcal{G}_{12}(u_1,h_1,V),$$

where

$$\begin{aligned}\mathcal{G}_{12}(u_1,h_1,V) &:= \mathbb{E}^{\mathbb{P}^1}[u_1(S_1)] + \sum_{K\in\mathcal{K}_2} f_2^{K,\text{bid/ask}}(V_K,\omega_K) + \sum_{K\in\mathcal{K}_2} V_K\mathcal{C}_2^{K,\text{mid}} \\ &\quad + \mathbb{E}^{m_0}\left[e^{-\sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+ - u_1(S_1) - h(S_1)(S_2-S_1)}\right].\end{aligned}$$

Vanishing the gradient with respect to u_1, we obtain the equation:

$$e^{-u_1(s_1)} = \frac{\mathbb{P}^1(s_1)}{I_u(h(s_1),V(\cdot),s_1)} \tag{4.4}$$

where we have set

$$I_u(\theta,V,s_1) := \int e^{-\sum_{K\in\mathcal{K}_2} V_K(s_2-K)^+ - \theta(s_2-s_1)}\mathbb{P}_0(s_1,ds_2) = e^{u_1(s_1)}\partial_{u_1(s_1)}\mathcal{G}_{12}(u_1,h_1,V),$$

Vanishing the gradient with respect to $h(s_1)$, we obtain the equation: $h(s_1) := \theta$ is the unique zero of

$$I_h(\theta,V,s_1) := 0, \tag{4.5}$$

where

$$I_h(\theta,V,s_1) := \int e^{-\sum_{K\in\mathcal{K}_2} V_K(s_2-K)^+ - \theta(s_2-s_1)}(s_2-s_1)m_0(s_1,ds_2) = e^{u_1(s_1)}\partial_{h_1(s_1)}\mathcal{G}_{12}(u_1,h_1,V).$$

In practice, this may be done thanks to a 1D Newton algorithm on the function $h_1(s_1) \mapsto \min_{h(s_1)}\mathcal{G}_{12}(u_1,h_1,V)$, see Subsections 3.3.3 and 3.3.5 in [8].

For use below, we also introduce for all $Q \in \mathcal{K}_2$:

$$I_Q(h(S_1),V,S_1) := \int (s_2-Q)^+ e^{-\sum_{K\in\mathcal{K}_2} V_K(s_2-K)^+ - \theta(s_2-s_1)} m_0(s_1,ds_2) = e^{u_1(s_1)}\partial_{V_Q}\mathcal{G}_{12}(u_1,h_1,V).$$

4.1. **Speed-up: Choice of a prior.** We take $m_0(ds_1,ds_2) = \mathbf{1}_{s_1\geq 0}\mathbb{P}_{\sigma_0}(ds_1)\mathbb{P}(ds_2|s_1)$ where $\mathbb{P}_{\sigma_0}(ds_1)$ is a normal density with volatility σ_0 and under $\mathbb{P}$:

$$S_2 = S_1 + \sigma(S_1)\sqrt{t_2-t_1}Z, \quad Z\in \mathrm{N}(0,1), \quad \sigma(S_1) = \sigma_0 S_1^\beta,$$

where σ_0 and β are two parameters. We choose σ_0 and β by minimizing the least-square problem:

$$\begin{aligned}&\inf_{\sigma_0,\beta}\sum_{K\in\mathcal{K}_2}\left(\mathbb{E}^{m_0}[(S_2-K)^+] - \mathcal{C}_2^{K,\text{mid}}\right)^2 \\ =\ &\inf_{\sigma_0,\beta}\sum_{K\in\mathcal{K}_2}\left(\mathbb{E}^{\mathbb{P}_{\sigma_0}}[\mathrm{B}(S_1,t_2-t_1,K,\sigma_0 S_1^\beta)] - \mathcal{C}_2^{K,\text{mid}}\right)^2,\end{aligned}$$

with

$$\mathrm{B}(s,t,K,\sigma) := \frac{1}{2}(s-K)\mathrm{erf}\left(\frac{K-s}{\sqrt{2}\sigma\sqrt{t}}\right) + \frac{\sigma\sqrt{t}e^{-\frac{(K-s)^2}{2\sigma^2 t}}}{\sqrt{2\pi}}.$$

As conditional on S_1, S_2 is normally-distribution, the integration over s_2 can be performed exactly in the definition of the functions I_u, I_h and I_Q, defined above, and they can be written in closed-form.

Remark 4.2 (Explicit formulas). For completeness, we give the formulas, obtained with Mathematica, that we use in our numerical implementation. Let $A := \frac{K_1 - s_1}{\sigma}$, $B := \frac{K_2 - s_1}{\sigma}$ and $\sigma := \sigma(s_1)\sqrt{t_2 - t_1}$. We have

$$\int_{K_1}^{K_2} e^{\alpha s_2}\mathbb{P}(ds_2|s_1) = \frac{1}{2}e^{\frac{\alpha^2\sigma^2}{2}+\alpha s_1}\left(\mathrm{erf}\left(\frac{B-\alpha\sigma}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{A-\alpha\sigma}{\sqrt{2}}\right)\right).$$

$$2\sqrt{2\pi}\int_{K_1}^{K_2} e^{\alpha s_2}(s_2 - s_1)\mathbb{P}(ds_2|s_1) = \sigma e^{\alpha s_1}\left(2e^{A\alpha\sigma - \frac{A^2}{2}} - \sqrt{2\pi}\alpha\sigma e^{\frac{\alpha^2\sigma^2}{2}}\mathrm{erf}\left(\frac{A-\alpha\sigma}{\sqrt{2}}\right)\right.$$
$$\left.+\sqrt{2\pi}\alpha\sigma e^{\frac{\alpha^2\sigma^2}{2}}\mathrm{erf}\left(\frac{B-\alpha\sigma}{\sqrt{2}}\right) - 2e^{\alpha B\sigma - \frac{B^2}{2}}\right).$$

$$2\sqrt{2\pi}\int_{K_1}^{K_2} e^{\alpha s_2}(s_2 - K)\mathbb{P}(ds_2|s_1) = e^{\alpha s_1}\left(2\sigma e^{A\alpha\sigma - \frac{A^2}{2}} - \sqrt{2\pi}e^{\frac{\alpha^2\sigma^2}{2}}\mathrm{erf}\left(\frac{A-\alpha\sigma}{\sqrt{2}}\right)\left(\alpha\sigma^2 - K + s_1\right)\right.$$
$$\left.+\sqrt{2\pi}e^{\frac{\alpha^2\sigma^2}{2}}\mathrm{erf}\left(\frac{B-\alpha\sigma}{\sqrt{2}}\right)\left(\alpha\sigma^2 - K + s_1\right) - 2\sigma e^{\alpha B\sigma - \frac{B^2}{2}}\right).$$

$$2\sqrt{2\pi}\int_{K_1}^{K_2} e^{\alpha(s_2 - s_1)}(s_2 - K)(s_2 - Q)\mathbb{P}(ds_2|s_1) = 2\sigma\left(e^{\alpha A\sigma - \frac{A^2}{2}}(\sigma(\alpha\sigma + A) - K - Q + 2s_1)\right.$$
$$\left.+e^{-\frac{1}{2}B(B-2\alpha\sigma)}(-\sigma(\alpha\sigma + B) + K + Q - 2s_1)\right)$$
$$+\sqrt{2\pi}e^{\frac{\alpha^2\sigma^2}{2}}\left(\mathrm{erf}\left(\frac{B-\alpha\sigma}{\sqrt{2}}\right) - \mathrm{erf}\left(\frac{A-\alpha\sigma}{\sqrt{2}}\right)\right)\left(\sigma^2 - \left(-\alpha\sigma^2 + K - s_1\right)\left(\alpha\sigma^2 - Q + s_1\right)\right).$$

The last formula is used for computing the hessian $\partial_V^2\mathcal{G}_{12}$.

Remark 4.3 (Other formulas). Note that we have

$$\mathbb{E}^{m_0}\left[e^{-\sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+ - h(S_1)(S_2-S_1) - u_1(S_1)}\right] = \mathbb{E}^{\mathbb{P}_{\sigma_0}}\left[I_u(h(S_1), V(\cdot), S_1)e^{-u_1(S_1)}\right].$$

and

$$\mathbb{E}^{\mathbb{P}_{\sigma_0}}\left[(S_2 - Q)^+ e^{-\sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+ - h(S_1)(S_2-S_1) - u_1(S_1)}\right] = \mathbb{E}^{\mathbb{P}_{\sigma_0}}\left[I_Q(h(S_1), V(\cdot), S_1)e^{-u_1(S_1)}\right].$$

4.2. **Sinkhorn's algorithm in a nutshell.** In order to be able to apply the next algorithm, we need to be able to do computations for each s_1. In order to do it in practice, we need to introduce an approximation $\mathbb{P}^1_{\mathcal{X}_1} \approx \mathbb{P}^1$ that is supported on a finite grid $\mathcal{X}_1 \subset \mathbb{R}_+$. As our goal is estimating call prices, that are naturally 1−Lipschitz, this approximation should be made in terms of Wasserstein distance $\mathcal{W}_1$. For the sake of simplicity, we approximate $\mathbb{P}^1$ by $\mathbb{P}^1_{\mathcal{X}_1} := \frac{1}{Z}\sum_{s_1\in\mathcal{X}_1}\frac{d\mathbb{P}^1}{dx}(s_1)\mathbf{1}_{s_1}$, where $Z := \sum_{s_1\in\mathcal{X}_1}\frac{d\mathbb{P}^1}{dx}(s_1)$ and $\mathcal{X}_1 := \{a + \frac{k}{n-1}(b-a) : 0 \le k \le n-1\}$, where $a < b$ and $n \ge 2$ are well chosen parameters in order to achieve convergence.

I Start with $h_1 := 0$, $u_1 := 0$ and $V_K := 0$ for all $K \in \mathcal{K}_2$.
II Projection on (u_1, h_1): Solve equations (4.4) and (4.5) for all $s_1 \in \mathcal{X}_1$.
III Solve the strictly convex smooth finite-dimensional unconstrained minimization over V:

$$\inf_{V_K\in\mathbb{R}} \qquad \mathcal{G}_{12}(u_1, h, V),$$

with

$$\begin{aligned}\mathcal{G}_{12}(u_1,h,V) &:= \mathbb{E}^{\mathbb{P}^1}[u_1]+\sum_{K\in\mathcal{K}_2} f_2^{K,\text{bid/ask}}(V_K,\omega_K)+\sum_{K\in\mathcal{K}_2} V_K\mathcal{C}_2^{K,\text{mid}}\\ &\quad +\mathbb{E}^{m_0}\left[e^{-\sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+-h(S_1)(S_2-S_1)-u_1(S_1)}\right].\end{aligned}$$

From Remark 4.3, this can be written exactly as

$$\begin{aligned}\mathcal{G}_{12}(u_1,h_1,V) &= \mathbb{E}^{\mathbb{P}^1}[u_1]+\sum_{K\in\mathcal{K}_2} f^{K,\text{bid/ask}}(V_K,\omega_K)+\sum_{K\in\mathcal{K}_2} V_K\mathcal{C}_2^{K,\text{mid}}\\ &\quad +\mathbb{E}^{\mathbb{P}_{\sigma_0}}[I_u(h(S_1),V,S_1)e^{-u_1(S_1)}].\end{aligned} \tag{4.6}$$

The gradients with respect to V can also be written as

$$\partial_{V_K} f^{K,\text{bid/ask}}(V_K,\omega_K)+\mathcal{C}_2^{K,\text{mid}}-\mathbb{E}^{\mathbb{P}_{\sigma_0}}[I_K(h(S_1),V,S_1)e^{-u_1(S_1)}]. \tag{4.7}$$

Similar formula for the hessian with respect to V.

IV Iterate steps (II)-(III) until convergence.

V Compute then the smile at t_2 for all $K\in\mathbb{R}_+$ (this defines the marginal $\mathbb{P}^2$):

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^*}[(S_2-K)^+] &= \mathbb{E}^{\mathbb{P}_{\sigma_0}}[I_K(h(S_1),V,S_1)e^{-u_1(S_1)}]\\ &\approx \mathbb{E}^{\mathbb{P}^1_{\mathcal{X}_1}}[I_K(h(S_1),V,S_1)e^{-u_1(S_1)}].\end{aligned}$$

4.3. **Convergence.** We define $0<K_1<...<K_k$ such that $\mathcal{K}_2:=\{K_i:1\le i\le k\}$. Furthermore, we abuse notation and denote $\mathbb{P}^1$ for $\mathbb{P}^1_{\mathcal{X}_1}$.

Definition 4.4. We say that $\left(\mathbb{P}^1,\left(\mathcal{C}_2^{K,\text{bid/ask}}\right)_{K\in\mathcal{K}_2}\right)$ is non-degenerate if up to denoting $K_0:=0$, and setting $\mathcal{C}_2^{0,\text{bid}}=\mathcal{C}_2^{0,\text{ask}}=\mathbb{E}^{\mathbb{P}^1}[S_1]$, we may find $\mathcal{C}\in\mathbb{R}^{k+1}$ such that for all $0\le i\le k$, we have $\mathcal{C}_2^{K_i,\text{bid}}\le\mathcal{C}_i\le\mathcal{C}_2^{K_i,\text{ask}}$, $(M_{\text{call}}^{-1}\mathcal{C})_i>0$, and if $i>0$ then $\mathcal{C}_i>\mathbb{E}^{\mathbb{P}^1}[(S_1-K_i)^+]$, where $M_{\text{call}}:=\left((K_{j+1}-K_i)^+\right)_{0\le i,j\le k}$, with the convention $(K_{k+1}-K_i)^+:=1$ for all i.

Theorem 4.5 (Convergence rate)**.** *The map $\mathcal{G}_{12}$ reaches a minimum $\mathcal{G}_{12}^*$ at some $x^*\in\mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}$ if and only if $\left(\mathbb{P}^1,\left(\mathcal{C}_2^{K,\text{bid/ask}}\right)_{K\in\mathcal{K}_2}\right)$ is non-degenerate.*

In this case, let $x_0=(u_0,h_0,V_0)\in\mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}$, and for $n\ge 0$, let the nth iteration of the well-defined martingale Sinkhorn algorithm:

$$\begin{aligned}x_{n+1/2} &:= \left(u_n,h_n,V_{n+1}:=\operatorname*{argmin}_{\psi}\mathcal{G}_{12}(u_n,h_n,\cdot)\right),\\ x_{n+1} &:= \left(u_{n+1}:=\operatorname*{argmin}_{u}\mathcal{G}_{12}(\cdot,\cdot,V_{n+1}),h_{n+1}:=\operatorname*{argmin}_{h}\mathcal{G}_{12}(\cdot,\cdot,V_{n+1}),V_{n+1}\right).\end{aligned}$$

Then we may find $0<\lambda<1$, and $M>0$ such that

$$\begin{aligned}\mathcal{G}_{12}(x_n)-\mathcal{G}_{12}^* &\le \lambda^n\big(\mathcal{G}_{12}(x_0)-\mathcal{G}_{12}^*\big),\\ |x_n-x^*|\le M\sqrt{\lambda}^n \quad &\text{and} \quad \big|\nabla\mathcal{G}_{12}(x_n)\big|\le M\sqrt{\lambda}^n,\end{aligned}$$

for all $n\ge 0$.

Proof. For the sake of simplicity of notations, we denote $\mathcal{G}_{12}$ by $\mathcal{G}$.

Step 1: We assume that $\left(\mathbb{P}^1, \left(\mathcal{C}_2^{K,\text{bid/ask}}\right)_{K\in\mathcal{K}_2}\right)$ is non-degenerate. Let $\mathcal{C} \in \mathbb{R}^{|\mathcal{K}_2|}$ be a valid call prices vector. Let us prove that $\mathcal{G}$ reaches a minimum at some $x^* \in \mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}$. First we prove that $\lim_{|x|\to\infty}\mathcal{G}(x) = \infty$. Let $(x_n)_{n\geq 0} \subset \mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}$ such that $|x_n| \longrightarrow \infty$. We assume for contradiction that up to replacing x_n by a subsequence, $\mathcal{G}(x_n)$ is bounded from above by $A > 0$. Then up to taking a subsequence of (x_n), we may assume that $\frac{x_n}{|x_n|}$ converges to some $x \in \mathcal{U} := \left\{x' \in \mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}\right\}$. Now let $\Delta := \left((\delta_{S_1=s_1})_{s_1\in\mathcal{X}_1}, (\delta_{S_1=x}(S_2-S_1))_{x\in\mathcal{X}_1}, ((S_2-K)^+)_{K\in\mathcal{K}_2}\right)$ so that for $x \in \mathbb{R}^{2|\mathcal{X}_1|+|\mathcal{K}_2|}$, we have

$$x\cdot\Delta \quad = \quad \sum_{s_1\in\mathcal{X}_1} x_{S_1} + x_{|\mathcal{X}_1|+S_1}(S_2-S_1) + \sum_{K\in\mathcal{K}_2} x_{2|\mathcal{X}_1|+K}(S_2-K)^+.$$

Notice that $\mathcal{C}$ is the subgradient of $V \longmapsto \sum_{K\in\mathcal{K}_2} f_2^{K,\text{bid/ask}}(V_K,\omega_K) + \sum_{K\in\mathcal{K}_2} V_K\mathcal{C}_2^{K,\text{mid}}$ at some point. Let $V^0 \in \mathbb{R}^{|\mathcal{K}_2|}$ be a point at which this gradient is reached. Then if we denote $b := \sum_{K\in\mathcal{K}_2} f_2^{K,\text{bid/ask}}(V_K^0,\omega_K) + \sum_{K\in\mathcal{K}_2} V_K^0\mathcal{C}_2^{K,\text{mid}}$ and $a := \left((\mathbb{P}^1[\{s_1\}])_{s_1\in\mathcal{X}_1}, 0, \mathcal{C}\right)$, we have

$$\mathcal{G}(x) \geq a\cdot x + b + \int e^{-x_n\cdot\Delta}dm_0. \tag{4.8}$$

Case 1: We may find $(s_1,s_2) \in \mathcal{X}_1\times\mathbb{R}_+$ such that $x\cdot\Delta(s_1,s_2) < 0$. As $x\cdot\Delta(s_1,\cdot)$ is affine by parts, we may find $\varepsilon > 0$ and an open interval $s_2 \in I \subset \mathbb{R}$ such that $x\cdot\Delta(s_1,\cdot) \leq -\varepsilon$ on I. Then for x' close enough to x, we have $x'\cdot\Delta(s_1,\cdot) \leq -\frac{1}{2}\varepsilon$ on I. Then by (4.8), for n large enough we have

$$\begin{aligned}\mathcal{G}(x_n) \quad &\geq \quad a\cdot x_n + b + \int e^{-x_n\cdot\Delta}dm_0\\ &\geq \quad a\cdot x_n + b + m_0[\{y_1\}\times I]e^{|x_n|\frac{1}{2}\varepsilon}.\end{aligned}$$

Therefore, by the fact that $m_0[\{s_1\}\times I] > 0$, we have that $\mathcal{G}(x_n)$ diverges to ∞ as $|x_n| \longrightarrow \infty$, a contradiction.

Case 2: $x\cdot\Delta \geq 0$ on $\mathcal{X}_1\times\mathbb{R}_+$. Then $\mathcal{G}(x_n) \geq a\cdot x_n + b = |x_n|a\cdot\frac{x_n}{|x_n|} + b$. As we assumed that $\mathcal{G}(x_n)$ is bounded and $\frac{x_n}{|x_n|}$ converges to x, we have

$$a\cdot x \leq 0. \tag{4.9}$$

We denote $(u,h,V) := x$ and we have $x\cdot\Delta := u + h^\otimes + \sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+$. Let $\psi := \sum_{K\in\mathcal{K}_2} V_K(S_2-K)^+$. We have $\psi \geq -u - h^\otimes$. Then if we denote f, the convex hull of ψ on $\mathbb{R}_+$, we have $\psi \geq f$ and for all $s_1 \in \mathcal{X}_1$, we have $f \geq -u(s_1) - h(s_1)\cdot(S_2-s_1)$. Therefore, $f \geq -u$ from last functional inequality computed in $S_2 = s_1$. By the fact that f is the convex hull of ψ, which is piecewise affine, f is also piecewise affine on the same intervals. Therefore, we may find $\lambda_i \geq 0$ for all $1 \leq i \leq k$ such that $f = f(0) + \nabla f(0) + \sum_{i=1}\lambda_i(S_2-K_i)^+$. Now let $p_i(S_1) := \mathbf{1}_{S_1\geq K_i}$, $p := \sum_{i=1}^k \lambda_i p_i$, and $\theta_i(S_1,S_2) := (S_2-K_i)^+ - (S_1-K_i)^+ - p_i(S_1)\cdot(S_2-S_1)$, we have $x\cdot\Delta = (u+f)\oplus(\psi-f) + (h+p)^\otimes + \theta$. Thus we have

$$a\cdot x \quad = \quad \mathbb{E}^{\mathbb{P}^1}[u+f] + \sum_{i=0}^k \mu_i\mathcal{C}_i + \sum_{i=1}^k \lambda_i\big(\mathcal{C}_i - \mathbb{E}^{\mathbb{P}^1}[(S_2-K_i)^+]\big),$$

where $(\psi - f)(S_2) = \sum_{i=0}^k \mu_i (S_2 - K_i)^+$. Notice that $(\psi - f) \geq 0$, therefore for all $0 \leq i \leq k$, $\gamma_i = (\psi - f)(K_{i+1}) = (M_{\text{call}}^t \mu)_i \geq 0$, then $\mu = (M_{\text{call}}^t)^{-1}\gamma$, and $\mu \cdot \mathcal{C} = (M_{\text{call}}^{-1}\mathcal{C}) \cdot \gamma$. Finally,

$$a \cdot x = \mathbb{E}^{\mathbb{P}^1}[u + f] + (M_{\text{call}}^{-1}\mathcal{C}) \cdot \gamma + \sum_{i=1}^k \lambda_i \big(\mathcal{C}_i - \mathbb{E}^{\mathbb{P}^1}[(S_2 - K_i)^+]\big).$$

By (4.9), $a \cdot x$ is non-positive, and the term on the right is non-negative by non-degeneracy, therefore $a \cdot x = 0$, and therefore $\lambda_1 = ... = \lambda_k = 0$, $\gamma = 0$, and $u + f = 0$. Therefore $u = -f$, $\psi = f$, and $f(s) = f(0) + \nabla f(0) \cdot s$. Therefore, $x \cdot \Delta = (h + p)^{\otimes} + \nabla f(0)(S_2 - S_1)$. Therefore, as $x \cdot \Delta \geq 0$, we have that $h + p + \nabla f(0) = 0$. Finally $x \cdot \Delta = 0$ on $\mathcal{X}_1 \times \mathbb{R}_+$, and finally $x = 0$, which is a contradiction as $x \in \mathcal{U}$.

We proved that $\lim_{|x| \to \infty} \mathcal{G}(x) = \infty$. As $\mathcal{G}$ is convex, it reaches a minimum at some $x^* \in \mathbb{R}^{2|\mathcal{X}_1| + |\mathcal{K}_2|}$.

Step 2: Now we assume that $\mathcal{G}$ reaches a minimum. Let us denote x^* this minimum and let $\mathbb{P}^*(ds_1, ds_2) = m_0(ds_1, ds_2) e^{-\sum_{K \in \mathcal{K}_2} V_K^*(s_2 - K)^+ - u_1^*(s_1) - h_1^*(s_1)(s_2 - s_1)}$. By Proposition 4.1, we have that $\mathbb{P}^* \in \widetilde{\mathcal{M}}_{12}^{\text{Markov}}(\mathbb{P}^1)$. Notice also that the measure $\mathbb{P}^*$ is equivalent to the measure m_0. Therefore, for all i the map θ_i is not equal to 0, $\mathbb{P}^*-$a.e. and therefore $\mathbb{E}^{\mathbb{P}^*}[\theta_i] > 0$. Finally we observe that if we denote $\mathcal{C}_i := \mathbb{E}^{\mathbb{P}^*}[(S_2 - K)^+]$, then we have $\mathbb{E}^{\mathbb{P}^*}[\theta_i] = \mathcal{C}_i - \mathbb{E}^{\mathbb{P}^1}[(S_1 - K)^+]$ from the martingale property of $\mathbb{P}^*$, and $\mathcal{C}_2^{K_i,\text{bid}} \leq \mathcal{C}_i \leq \mathcal{C}_2^{K_i,\text{ask}}$ as $\mathbb{P}^* \in \widetilde{\mathcal{M}}_{12}^{\text{Markov}}(\mathbb{P}^1)$. Now for $1 \leq i \leq k$, let f_i the piecewise affine map such that f is zero on $[0, K_{i-1}]$, with $f(K_i) = 1$, affine on $[K_{i-1}, K_i]$, $[K_i, K_{i+1}]$, and $[K_{i+1}, \infty)$, if $i \neq k$, and f_k is constant equal to 1 on $[K_k, \infty]$. We observe that for all i, f_i is non-negative and non-zero $\mathbb{P}^*-$a.e. Furthermore, $0 < \mathbb{E}^{\mathbb{P}^*}[f_i] = (M_{\text{call}}^{-1}\mathcal{C})_i$. We proved that $\left(\mathbb{P}^1, \left(\mathcal{C}_2^{K,\text{bid/ask}}\right)_{K \in \mathcal{K}_2}\right)$ is non-degenerate.

Step 3: The convergence result stems from an indirect application of Theorem 5.2 in [3]. By a direct application of this theorem we get that

$$\mathcal{G}(x_k) - \mathcal{G}(x^*) \leq \left(1 - \tfrac{\sigma}{\min(L_1, L_2)}\right)^{n-1} \big(\mathcal{G}(x_0) - \mathcal{G}(x^*)\big), \tag{4.10}$$

with L_1 (resp. L_2) is the Lipschitz constant of the $V-$gradient (resp. $(u, h)-$gradient) of $\mathcal{G}$, and σ is the strong convexity parameter of $\mathcal{G}$. Furthermore, the strong convexity gives that

$$|x_k - x^*| \leq \sqrt{\frac{2}{\sigma}} \big(\mathcal{G}(x_k) - \mathcal{G}(x^*)\big)^{\frac{1}{2}}. \tag{4.11}$$

Finally, by definition of L_1 and L_2, we have

$$|\nabla \mathcal{G}(x_k)| \leq (L_1 + L_2)|x_k - x^*|. \tag{4.12}$$

However the gradient $\nabla \mathcal{G}$ is locally but not globally Lipschitz, nor $\mathcal{G}$ strongly convex. Therefore we need to refine the theorem by looking carefully at where these constants are used in its proof.

Step 4: The constant L_1 is used for Lemma 5.1 in [3]. We need for all $k \geq 0$ to have $\mathcal{G}(x_k) - \mathcal{G}(x_{k+1/2}) \geq \frac{1}{2L_1}|\nabla \mathcal{G}(x_k)|^2$. We want to find $C, L > 0$ such that $\mathcal{G}(x_k) - \mathcal{G}\big(x_k - C\nabla \mathcal{G}(x_k)\big) \geq \frac{1}{2L}|\nabla \mathcal{G}(x_k)|$, then L may be use to replace L_1 in the final step of the proof of Lemma 5.1 in [3]. By the fact that $\lim_{|x| \to \infty} \mathcal{G}(x) = \infty$, the set $\mathfrak{C}(x_0) := \{x \in \mathbb{R}^{2|\mathcal{X}_1| + |\mathcal{K}_2|} : \mathcal{G}(x) \leq \mathcal{G}(x_0)\}$ is compact. Then $\nabla \mathcal{G}$ is bounded on $\mathfrak{C}(x_0)$. Therefore we may find $M_1 > 0$ such that for all k, we have $|\nabla \mathcal{G}(x_k) \cdot$

$\Delta| \leq M_1(1+|S_2|)$. Furthermore, let $F(C) := \sup_{(u,x)\in\mathcal{U}\times\mathfrak{C}(x_0)} \int_0^1 \int (u\cdot\Delta)e^{-x\cdot\Delta}e^{tCM_1(1+|S_2|)}dm_0dt$. We have

$$\begin{aligned}
\mathcal{G}(x_k) - \mathcal{G}\big(x_k - C\nabla\mathcal{G}(x_k), h_k\big) &= \nabla\mathcal{G}(x_k)\cdot\big(-C\nabla\mathcal{G}(x_k)\big) \\
&\quad -C^2\int_0^1 (1-t)D^2\mathcal{G}\big(x_k - tC\nabla\mathcal{G}(x_k)\big)\big(\nabla\mathcal{G}(x_k)\big)^2 dt \\
&= C|\nabla\mathcal{G}(x_k)|^2 \\
&\quad -C^2\int_0^1 (1-t)\int(\nabla\mathcal{G}(x_k)\cdot\Delta)^2 e^{-x_k\cdot\Delta}e^{tC\nabla\mathcal{G}(x_k)\cdot\Delta}dm_0dt \\
&\geq C|\nabla\mathcal{G}(x_k)|^2 - C^2|\nabla\mathcal{G}(x_k)|^2F(C) \\
&= \big(C - C^2F(C)\big)|\nabla\mathcal{G}(x_k)|^2.
\end{aligned}$$

As F is non-decreasing finite, then when $C \longrightarrow 0$ we have $\frac{C-C^2F(C)}{C} \longrightarrow 1$. Then for C small enough, let $L := \frac{1}{2\big(C-C^2F(C)\big)} > 0$. We get

$$\mathcal{G}(x_k) - \mathcal{G}\big(x_k - C\nabla\mathcal{G}(x_k)\big) \geq \frac{1}{2L}|\nabla\mathcal{G}(x_k)|^2.$$

Step 5: The constant σ is used to get the result from (3.21) in [3]. Then we just need the inequality

$$\mathcal{G}(y) \geq \mathcal{G}(x) + \nabla\mathcal{G}(x)\cdot(y-x) + \frac{\sigma}{2}|y-x|^2, \tag{4.13}$$

to hold for some $y = x^*$ and $x = x_k$ for all $k \geq 0$. Now we give a lower bound for σ. The map $(u,x) \longmapsto D^2\mathcal{G}(x)u^2 = \int(u\cdot\Delta)^2e^{-x\cdot\Delta}dm_0 > 0$ is continuous on $\mathcal{U}\times\mathfrak{C}(x_0)$ compact, therefore it has a lower bound $\sigma > 0$. This constant also works for (4.11). Similar, $\sup_{(u,x)\in\mathcal{U}\times\mathfrak{C}(x_0)} D^2\mathcal{G}(x)u^2$ may replace $L_1 + L_2$ from (4.12).

Step 6: Finally, as we focus on the L_1 optimization phase, we may replace $n-1$ by n in the convergence formula (4.10), see the proof of Theorem 5.2 in [3].

Now the existence of $M > 0$ stems from the facts that $\mathcal{G}(x_k) - \mathcal{G}(x^*) \geq \frac{1}{2}\sigma|x_k - x^*|^2$, and $|\nabla\mathcal{G}(x_k)| \leq L|x_k - x^*|$.

The result is proved. □

Remark 4.6. Notice that

$$\begin{aligned}
\nabla\mathcal{G} &= \sum_{s_1\in\mathcal{X}_1}\big(\mathbb{P}^1_{\mathcal{X}_1}[\{s_1\}] - \mathbb{P}\circ(S_1)^{-1}[\{s_1\}]\big)\,e_{s_1} + \sum_{s_1\in\mathcal{X}_1}\big(\mathbb{E}^{\mathbb{P}}[S_2 - s_1, S_1 = s_1]\big)\,e_{|\mathcal{X}_1|+s_1} \\
&\quad + \sum_{i=1}^{k}\big(C_K - \mathbb{E}^{\mathbb{P}}[(S_2 - K_i)^+]\big)\,e_{2|\mathcal{X}_1|+i},
\end{aligned}$$

where $C_K \in [\mathcal{C}_2^{K,\text{bid}}, \mathcal{C}_2^{K,\text{ask}}]$, and $(e_i)_{1\leq i\leq 2|\mathcal{X}_1|+|\mathcal{K}_2|}$ is the canonical basis. Therefore, this gradient is a crucial estimate of the mismatch of $\mathbb{P}$ in terms of first marginal, martingale property, and correctness of the call prices it gives.

Remark 4.7. Minimizing $\mathcal{G}$ using the Sinkhorn algorithm is classical but we may obtain better stability and speed of convergence by using an implied Newton minimization algorithm, see

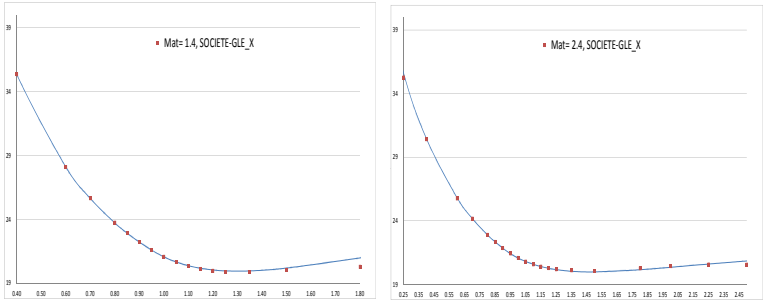

FIGURE 3. SOCIETE-GENERALE.

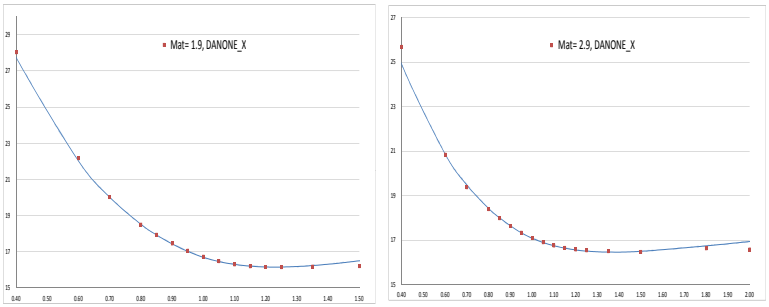

FIGURE 4. DANONE.

3.3.5. in [8]. This algorithm consists in applying a truncated Newton algorithm on $\widetilde{\mathcal{G}}(V) := \min_{u_1,h_1} \mathcal{G}(u_1, h_1, V)$ which is also strictly convex and smooth, see Proposition 3.2 in [8]. This algorithm should converge faster as we need anyways to use a Newton algorithm of the same dimension $|\mathcal{K}_2|$ for the partial minimization in V during phase (III) of the Sinkhorn algorithm.

Remark 4.8. Even though the criterion from Definition 4.4 may not be easy to compute, trying to solve the entropic minimization reveals if a solution exists as otherwise the map $\mathcal{G}$ diverges to $-\infty$. In this case there is an arbitrage between the call prices and $\mathbb{P}^1$.

Remark 4.9. As we solve the problem building $\mathbb{P}^{i+1}$ after having built $\mathbb{P}^i$, we may encounter a situation in which the probability $\mathbb{P}^1$ may not be compatible with the call prices available in the intervals $[\mathcal{C}_2^{K,\text{bid}}, \mathcal{C}_2^{K,\text{ask}}]$. In this case it is necessary to solve the problem globally by including all the times in the value function and applying a wider Sinkhorn algorithm, minimizing on each time one after the other like it is done for example in [4]. We have not encountered this situation in our numerical experiments using real market datas.

4.4. **Numerical examples.** Below, we list some numerical examples involving numerous equity stocks/indices with various liquidity/maturities: Socíété Générale, Danone, Apple, SP500.

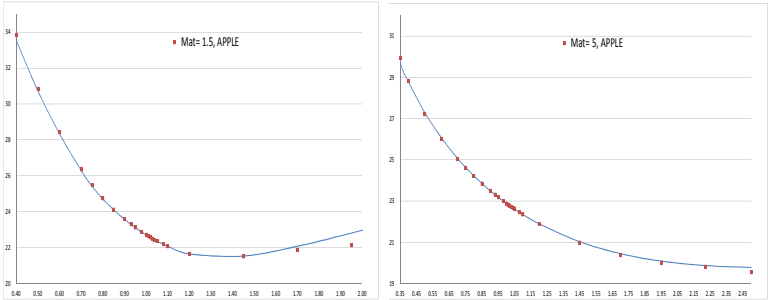

FIGURE 5. APPLE.

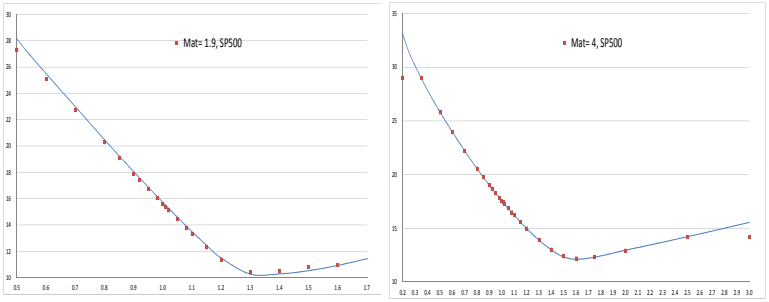

FIGURE 6. SP500.

REFERENCES

[1] Jesper Andreasen and Brian Huge. Expanded forward volatility. *Risk*, 26(1):101, 2013.

[2] Marco Avellaneda, Robert Buff, Craig Friedman, Nicolas Grandechamp, Lukasz Kruk, and Joshua Newman. Weighted monte carlo: a new technique for calibrating asset-pricing models. *International Journal of Theoretical and Applied Finance*, 4(01):91–119, 2001.

[3] Amir Beck and Luba Tetruashvili. On the convergence of block coordinate descent type methods. *SIAM journal on Optimization*, 23(4):2037–2060, 2013.

[4] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.

[5] Peter Carr and Dilip B Madan. A note on sufficient conditions for no arbitrage. *Finance Research Letters*, 2(3):125–130, 2005.

[6] Laurent Cousot. Conditions on option prices for absence of arbitrage and exact calibration. *Journal of Banking & Finance*, 31(11):3377–3397, 2007.

[7] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*, pages 2292–2300, 2013.

[8] Hadrien De March. Entropic resolution for multi-dimensional optimal transport. *arXiv preprint arXiv:1812.11104*, 2018.

[9] Bruno Dupire. Pricing and hedging with smiles. *Mathematics of derivative securities*, 1(1):103–111, 1997.

[10] Peter K Friz, Jim Gatheral, Archil Gulisashvili, Antoine Jacquier, and Josef Teichmann. *Large deviations and asymptotic methods in finance*, volume 110. Springer, 2015.

[11] Jim Gatheral and Antoine Jacquier. Convergence of heston to svi. *Quantitative Finance*, 11(8):1129–1132, 2011.

[12] Jim Gatheral and Antoine Jacquier. Arbitrage-free svi volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.

[13] Gaoyue Guo and Jan Obloj. Computational methods for martingale optimal transport problems. *arXiv preprint arXiv:1710.07911*, 2017.

[14] Pierre Henry-Labordère. *Analysis, geometry, and modeling in finance: Advanced methods in option pricing.* Chapman and Hall/CRC, 2008.

[15] Pierre Henry-Labordere. Calibration of local stochastic volatility models to market smiles: A monte-carlo approach. 2009.

[16] Pierre Henry-Labordère. Automated option pricing: Numerical methods. *International Journal of Theoretical and Applied Finance*, 16(08):1350042, 2013.

[17] Pierre Henry-Labordere and Nizar Touzi. An explicit martingale version of brenier's theorem. 2013.

[18] Alex Lipton and Artur Sepp. Filling the gaps. 2011.

[19] Alexander Lipton. Masterclass with deutsche bank. the vol smile problem. *RISK-LONDON-RISK MAGAZINE LIMITED-*, 15(2):61–66, 2002.

[20] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport. Technical report, 2017.

CMAP, École Polytechnique

E-mail address: `hadrien.de-march@polytechnique.org`

Société Générale, Global Market Quantitative Research

E-mail address: `pierre.henry-labordere@sgcib.com`