Profitable forecast of prices of stock options on real market data via the solution of an ill-posed problem for the Black-Scholes equation

Michael V. Klibanov∗ and Andrey V. Kuzhuget◦

∗Department of Mathematics & Statistics
University of North Carolina at Charlotte
Charlotte, NC 28223.
◦ Morgan Stanley, New York, NY 10036.
Emails: mklibanv@uncc.edu and akuzhuget@gmail.com

Abstract
A new mathematical model for the Black-Scholes equation is proposed to forecast option prices. This model includes new interval for the price of the underlying stock as well as new initial and boundary conditions. Conventional notions of maturity time and strike prices are not used. The Black-Scholes equation is solved as a parabolic equation with the reversed time, which is an ill-posed problem. Thus, a regularization method is used to solve it. This idea is verified on real market data for twenty liquid options. A trading strategy is proposed. This strategy indicates that our method is profitable on at least those twenty options. We conjecture that our method might lead to significant profits of those financial institutions which trade large amounts of options. We caution, however, that detailed further studies are necessary to verify this conjecture.

Key Words: Black-Scholes equation, new mathematical model, new initial and boundary conditions, testing on real market data, parabolic equation with the reversed time, Ill-Posed problem, regularization method

1 Introduction
The Black-Scholes equation is solved forwards in time to forecast prices of stock options. Since this is an ill-posed problem, a regularization method is used. Uniqueness, stability and convergence theorems for this method are formulated. For each individual option, we use its past history as an input data. We propose a new mathematical model for the interval of prices of the underlying stock, as well as for initial and boundary conditions on this interval for the Black-Scholes equation. The conventional notions of strike prices and maturity time are not used. Our model does not make distinctions between put and call options. Based on our results for market data, we conjecture that our methodology
might result in significant profits if trading a large number of options, as it is done in some large hedge funds. We caution, however, that additional studies of a substantial number of options are necessary to verify this conjecture.

To verify the validity of our model, we use the real market data for twenty liquid traded options. We believe that this is the best possible way for the verification. We select options randomly among liquid ones. Those liquid options were taken selected at \url{http://finance.yahoo.com/options/lists/}. Market prices, implied volatility and prices of the underlying stock for our selected options were taken from the Bloomberg terminal \url{http://www.bloomberg.com}. The single condition we impose when selecting these options, is that these options should be daily traded. Based on our technique, we propose here a certain trading strategy. This strategy shows that we are profitable on seventeen (17) out of twenty (20) options.

Our mathematical model allows one to forecast option prices of daily traded options from the current time event to two next ones, i.e. for a short time period. As to the time distance between two events, it depends on time units which one is using. In our calculations one time unit is one trading day. Thus, we forecast prices for “tomorrow” and the “day after tomorrow” having knowledge of past prices of “the day before yesterday”, “yesterday” and “today”. For our convenience, we use only last prices in each particular day, i.e. prices which were in place just before the closure of the market. But one unit can be one minute, one hour, one week, etc., and also last prices can be replaced with prices at other moments of time. An optimization of the latter is a technical issue which can be addressed later.

In the conventional approach to the Black-Scholes equation, one solves this equation backwards in time \( t \in (0,T) \), having an initial condition at the maturity time \( t = T \). The initial condition at \( t = T \) includes the strike price \( K \), see Hull (2000) and Wilmott, Howison and Dewyne (1997). However, there is a major drawback in this approach. Indeed, \( T \) is usually a few months. But it is obviously impossible to forecast the value of the volatility for such a long time period with a reasonable accuracy. On the other hand, the solution of the Black-Scholes equation critically depends on the volatility coefficient.

The above led us to believe that it is more natural to use the Black-Scholes equation to forecast prices of stock options for short time periods. However, the first major obstacle in this direction is that the problem of solving the Black-Scholes equation forwards in time is ill-posed. Namely, its solution usually does not exists and even if it exists, then it is unstable with respect to the input data. Therefore, a regularization method should be used. The second major obstacle is that some crucial input parameters for solving that equation forwards in time are unknown. These parameters are: the interval of prices of the underlying stock, the initial condition, two boundary conditions and the volatility coefficient.

We positively address here the following two questions:

**Question 1.** Is it possible to forecast the option price for a rather short time period using the Black-Scholes equation and market data?

**Question 2.** If this is possible, then can one be profitable using this forecast and a trading strategy?

There are four main questions which need to be addressed to answer the first question:

1. What is the interval for the prices of the underlying stock?

2. What are boundary and initial conditions on this interval?
3. What are the values of the volatility coefficient in the future?

4. How to solve the Black-Scholes equation forwards in time?

We address questions 1-3 in our new mathematical model. To address the fourth question, we use a regularization algorithm which was developed in Klibanov (2014). We formulate theorems about stability and convergence of this method. They were proven in Klibanov (2014) for a general parabolic equation of the second order. The key method of proofs in these references is the method of Carleman estimates.

We now briefly explain why the solution of the Black-Scholes equation in the forward direction of time is an ill-posed problem. Let \( \tau = \text{const.} > 0 \) and the function \( f \in L_2(0, \pi) \). Here is a well known example demonstrating the ill-posedness of problems for parabolic equations with the reversed time. Consider the following problem for the heat equation with the reversed time,

\[
\begin{align*}
&u_t + u_{xx} = 0, (x, t) \in (0, \pi) \times (0, \tau), \\
&u(x, 0) = f(x), \\
&u(0, t) = u(\pi, t) = 0.
\end{align*}
\]

It is well known that this problem has no more than one solution, see, e.g. Klibanov (2014) and Lavrentiev, Romanov and Shishatskii (1986). The unique solution of this problem, if it exists, is

\[
\begin{align*}
u(x, t) &= \sum_{n=1}^{\infty} f_n \sin(nx) e^{n^2 t},
\end{align*}
\]

where \( \{f_n\}_{n=1}^{\infty} \) are Fourier coefficients of the function \( f(x) \) with respect to \( \sin(nx) \).

Consider the \( L_2(0, \pi) \) norm of the function \( u(x, t) \),

\[
\|u(x, t)\|^2_{L_2(0, \pi)} = \sum_{n=1}^{\infty} f_n^2 e^{2n^2 t}.
\] (1.1)

Hence, if the solution of this problem exists, then squares of Fourier coefficients \( f_n^2 \) must decay exponentially with respect to \( n \). Hence, the solution of this problem exists for only a narrow set of functions \( f \). Also, it is clear from (1.1) that even if this series is truncated, still small fluctuations of \( f_n^2 \) can lead large variations of the function \( u \). This manifests the instability of the above problem. Furthermore, it follows from (1.1) that the larger \( \tau \) is, the more unstable this problem is. Hence, in order to obtain a rather good accuracy, any regularization method should work only on a short time interval. Thus, a rather accurate forecast of option prices via the Black-Scholes equation might occur only for a rather short time period. The latter is exactly what we are doing here.

In section 2 we present our mathematical model. In section 3 we describe our numerical method. In section 4 we show our results for market data. We summarize our results in section 5.
2 The new mathematical model

Let \( s \) be the stock price, \( t \) be time, \( \sigma (t) \) be the volatility of the option. In our particular case we use only the implied volatility listed on the market data of [http://www.bloomberg.com](http://www.bloomberg.com). However, more sophisticated models for the volatility can also be used in our model. Let \( \tau > 0 \) be the unit of time for which we want to forecast the option price. In our particular case \( \tau = 1/255 \) days. Hence, our model can work with the case when \( \tau \) is any unit of time: one hour, one minute, etc. In any case we assume below that \( \tau \in (0,1/4) \). Let “today” be \( t = 0 \), “tomorrow” be \( t = \tau \), the day after “tomorrow” be \( t = 2\tau \), “yesterday” be \( t = -\tau \) and “the day before yesterday” be \( t = -2\tau \).

We forecast only the last price of the option, i.e. the price on which one item of that option was either bought or sold in the last trading event of a trading day for this particular option. Let \( u_b (t) \) and \( u_a (t) \) be respectively the bid and ask prices of the option at the moment of time \( t \). Let \( s_b (t) \) and \( s_a (t) \) be respectively the bid and ask prices of the stock at the moment of time \( t \). It is well known that \( u_b (t) < u_a (t) \) and that \( s_b (t) < s_a (t) \).

Thus, the market data which we need for our model are listed in Table 1. These data are available at [http://www.bloomberg.com](http://www.bloomberg.com).

| \( t \) | \( u_b (t) \) | \( u_a (t) \) | \( \sigma (t) \) | \( s_b (t) \) | \( s_a (t) \) |
|-------|-------------|-------------|-------------|-------------|-------------|
| \( t = -2\tau, -\tau, 0 \) | \( t = -2\tau, -\tau, 0 \) | \( t = -2\tau, -\tau, 0 \) | \( t = 0 \) | \( t = 0 \) |

First, having discrete values of functions \( u_b (t), u_a (t), \sigma (t) \) for three moments of time listed in this table, we interpolate these functions between these three points using the standard quadratic interpolation. Thus, we obtain approximate values of these functions for \( t \in (-2\tau, 0) \) as quadratic polynomials. Next, we extrapolate functions \( u_b (t), u_a (t) \) for \( t \in (0, 2\tau) \) as those quadratic polynomials. Since \( \tau \) is small, then it is reasonable to assume that functions \( u_b (t), u_a (t), \sigma (t) \) are approximated rather well on the time interval \( t \in (-2\tau, 2\tau) \). Thus, we obtain functions

\[
\begin{align*}
  u_b (t), u_a (t), \sigma (t) & \quad \text{for } t \in (0, 2\tau).
\end{align*}
\]

Let \( u = u (s,t) \) be the price of one item of the stock option. Denote \( s_b = s_b (0), s_a = s_a (0) \). Then \( s_b < s_a \). Consider the numbers \( u_b = u (s_b, 0) \) and \( u_a = u (s_a, 0) \). Let

\[
  f (s) = \frac{u_b - u_a}{s_b - s_a} + \frac{u_a s_b - u_b s_a}{s_b - s_a}
\]

be the linear interpolation between \( u_b \) and \( u_a \) on the interval \( s \in (s_b, s_a) \). Hence, \( f (s_b) = u_b, f (s_a) = u_a \).

The simplest form of the Black-Scholes equation is

\[
  Lu := u_t + \frac{\sigma^2 (t)}{2} s^2 u_{ss} = 0.
\]
where $L$ is the partial differential operator of the Black-Scholes equation, see Hull (2000) and Wilmott, Howison and Dewyne (1997). We solve this equation on the interval of stock prices $s \in (s_b, s_a)$ and for times $t \in (0, 2\tau)$. Thus, we impose the following initial condition at $t = 0$ and boundary conditions at $s = s_b, s_a$:

$$u(s, 0) = f(s), s \in (s_b, s_a),$$  \hspace{1cm} (2.3)

$$u(s_b, t) = u_b(t), u(s_a, t) = u_a(t), \text{ for } t \in (0, 2\tau).$$  \hspace{1cm} (2.4)

Denote $Q_{2\tau} = \{(s, t) : s \in (s_b, s_a), t \in (0, 2\tau)\}$. In this paper, we computationally solve the following problem:

**Problem.** For $(s, t) \in Q_{2\tau}$ find the solution $u(s, t)$ of equation (2.2) satisfying the initial condition (2.3) and boundary conditions (2.4).

This problem as well as the above interpolations and extrapolations form our new mathematical model. Theorem 1 claims uniqueness of the solution of this problem. This theorem follows immediately from Klibanov (2014) and Lavrentiev, Romanov and Shishatskii (1986). Below $H^{2,1}(Q_{2\tau})$ and $H^2(Q_{2\tau})$ are standard Sobolev spaces of real valued functions.

**Theorem 1.** The problem (2.2)-(2.4) has at most one solution $u \in H^{2,1}(Q_{2\tau})$.

### 3 Numerical method for the problem (2.2)-(2.4)

As it was pointed out in Introduction, the problem (2.2)-(2.4) is ill-posed, since we are trying to solve the Black-Scholes equation forwards rather than backwards in time. The ill-posedness means here that the existence of the solution is not guaranteed. In addition, the solution, even if it exists, is unstable with respect to small fluctuations of initial and boundary conditions, see, e.g. the book of Tikhonov, Goncharsky, Stepanov, and Yagola (1995) for the theory of Ill-Posed Problems. Therefore, we use the regularization method of section 5 of Klibanov (2014). In simple terms, we find such an approximate solution of this Problem, which satisfies conditions (2.2)-(2.4) in the best way in the least squares sense.

Consider the following function $F(s, t)$

$$F(s, t) = \frac{u_b(t) - u_a(t)}{s_b - s_a} \cdot s + \frac{u_a(t) s_b - u_b(t) s_a}{s_b - s_a}. \hspace{1cm} (3.1)$$

Then $F \in H^2(Q_{2\tau})$. It follows from (2.1), (2.3), (2.4) and (3.1) that

$$F(s, 0) = f(s), F(s_b, t) = u_b(t), F(s_a, t) = u_a(t).$$  \hspace{1cm} (3.2)

Following section 5 of Klibanov (2014), consider the following Tikhonov-like functional

$$J_\alpha(u) = \int_{Q_{2\tau}} (Lu)^2 \, dx \, dt + \alpha \|u - F\|^2_{H^2(Q_{2\tau})}, \hspace{1cm} (3.3)$$

where $\alpha \in (0, 1)$ is the regularization parameter. Note that in the conventional case of linear ill-posed problems Tikhonov functional is generated by a bounded linear operator, see, e.g. Ivanov, Vasin and Tanana (2002). However, in our case $L$ is an unbounded
differential operator $L : H^{2,1}(Q_{2r}) \to L_2(Q_{2r})$, in which case $H^{2,1}(Q_{2r})$ is considered as a dense linear set in the space $L_2(Q_{2r})$. We consider the following minimization problem:

**Minimization Problem.** Minimize the functional $J_\alpha(u)$ in (3.3), subject to the initial and boundary conditions (2.3), (2.4).

To solve this minimization problem computationally, we have written partial derivatives in (3.3) via finite differences. In particular, we have obtained a finite difference grid covering the rectangle $Q_{2r}$. Next, we have minimized $J_\alpha(u)$ with respect to the values of the function $u(s,t)$ at grid points, using the conjugate gradient method. The starting point of this method was $u \equiv 0$.

The existence and uniqueness of the minimizer of functional (3.3) supplied by boundary conditions (2.3), (2.4) is ensured by Theorem 2. This theorem follows immediately from Theorem 5.3 of Klibanov (2014), where the general parabolic equation of the second order was considered.

**Theorem 2.** In (3.3), let $F$ be the function defined in (3.2). Then for any $\alpha \in (0,1)$ there exists a unique minimizer $u_\alpha \in H^2(Q_{2r})$ of the functional $J_\alpha(u)$ satisfying conditions (2.3), (2.4). Furthermore, there exists a constant $C = C(Q_{2r}, \sigma) > 0$ depending only on listed parameters such that

$$\|u_\alpha\|_{H^2(Q_{2r})} \leq \frac{C}{\sqrt{\alpha}} \|F\|_{H^2(Q_{2r})} .$$

To formulate the convergence result for minimizers $u_\alpha$, we need to assume that there exists the “ideal” exact solution of our ill-posed problem, i.e. solution with the ideal noiseless data. Such an assumption is one of most important building blocks of the Tikhonov regularization theory, see, e.g. Tikhonov, Goncharsky, Stepanov, and Yagola (1995). Thus, we assume that there exists the exact solution $u^*(s,t) \in H^2(Q_{2r})$ of equation (2.2) with exact initial condition $f^*(s) \in H^2(s_b, s_a)$ in (2.3) and exact boundary conditions $u^*_b(t), u^*_a(t) \in H^2(0, 2\tau)$. We also assume that there exists a function $F^* \in H^2(Q_{2r})$ such that

$$F^*(s,0) = f^*(s), F^*(s_b,t) = u^*_b(t), F^*(s_a,t) = u^*_a(t) ,$$

which is similar with (3.2). Let $\delta \in (0,1)$ be a sufficiently small number. We assume that

$$\|F - F^*\|_{H^2(Q_{2r})} \leq \delta .$$

Hence, it follows from (3.2), (3.4) and (3.5) that the number $\delta$ characterizes the level of the error in our data $f(s), u_b(t), u_a(t)$ as compared with the exact data $f^*(s), u^*_b(t), u^*_a(t)$.

In Theorem 3 we estimate the convergence rate of our minimizers $u_\alpha$ to the exact solution $u^*$, assuming that $\delta \to 0$. Theorem 3 follows immediately from Theorem 5.4 of Klibanov (2014), which was proven for a general parabolic operator of the second order. As it is always done in the regularization theory, we choose in Theorem 3 the regularization parameter $\alpha = \alpha(\delta)$ depending on the level of the error in the data. Note that assumption (3.5) of the small perturbation in the data seems to be close to the reality. Indeed, although we do not assume that the function $f^*(s)$ is linear and functions $u^*_b(t), u^*_a(t)$ are quadratic ones, as our functions $f(s), u_b(t), u_a(t)$ are, still since intervals $(s_b, s_a)$ and $(0, 2\tau)$ are small, then it is reasonable to assume, as we do, that the function $f(s)$ is a linear and functions $u_b(t), u_a(t)$ are quadratic.
Theorem 3. Assume that conditions (3.4) and (3.5) hold. Let \( a \in (0, 1) \) be a number, the function \( \sigma (t) \in C^1 [0, a] \) and there exist constants \( \sigma_0, \sigma_1 > 0, \sigma_0 < \sigma_1 \) such that \( \sigma (t) \in [\sigma_0, \sigma_1] \) for \( t \in [0, a] \). Choose the regularization parameter \( \alpha = \alpha (\delta) = \delta^{2\beta} \), where \( \beta = \text{const.} \in (0, 1) \). Then there exists a sufficiently small number \( \tau_0 = \tau_0 (\sigma_0, \sigma_1, \|\sigma\|_{C^1[0, a]} ) \in (0, a) \) depending only on listed parameters such that if \( \tau \in (0, \tau_0) \), then the following convergence rates hold

\[
\| \partial_s u_{\alpha}(\delta) - \partial_s u^* \|_{L^2(Q_+) \cap \Gamma^0(Q_+)} + \| u_{\alpha}(\delta) - u^* \|_{L^2(Q_+)} \leq B \left( 1 + \| u^* \|_{H^2(Q_{2\tau})} \right) \delta^\gamma, \tag{3.6}
\]

\[
\| u_{\alpha}(\delta) (s, 2\tau) - u^* (s, 2\tau) \|_{L^2(s_\delta, s_a)} \leq \frac{B}{\sqrt{\ln (\delta^{-1})}} \left( 1 + \| u^* \|_{H^2(Q_{2\tau})} \right), \tag{3.7}
\]

where the number \( \gamma = (\beta \ln 2) / \ln 4 \) and the constant \( B = B (a, Q_{2\tau}, \sigma) > 0 \) depends only on listed parameters.

It is clear from this theorem that the accuracy estimate (3.6) in a smaller domain \( Q_{\tau} \subset Q_{2\tau} \) is of the Hölder type. On the other hand the estimate (3.7) on the upper side of the rectangle \( Q_{2\tau} \) is of the logarithmic type. Clearly, the accuracy guaranteed by (3.6) is better than the accuracy guaranteed by (3.7). Hence, one can expect to obtain more accurate computational results in \( Q_{\tau} \), as compared with those in \( Q_{2\tau} \). This is exactly the reason why we have chosen to solve the problem (2.2)-(2.4) in the larger domain \( Q_{2\tau} \). It is also clear from Theorem 3 why we have chosen the number \( \tau > 0 \) to be sufficiently small.

4 Results for the market data

In our computations we have always used \( \tau = 1/255 \), as mentioned in section 2. In other words, we forecasted option prices for two days ahead, i.e. for \( t \in (0, 2\tau) \). To make sure that we obtain accurate results at least for the computationally simulated data, we have solved first the above Minimization Problem for synthetic data. To simulate these data computationally, we have solved equation (2.2) downwards in time with boundary conditions (2.4) and the initial condition at \( t = 2\tau \). This way we have obtained the function \( u (s, 0) = f_{\text{sim}} (s) \). Next we have solved equation (2.2) with boundary conditions (2.4) and the initial condition \( f_{\text{sim}} (s) \) upwards in time by the above described method. Next, we have compared the resulting solution with the computationally simulated one. Results were quite accurate ones. In this study of computationally simulated data we have found that \( \alpha = 0.01 \) is the optimal value of the regularization parameter and we have used \( \alpha = 0.01 \) in all follow up computations.

Next, we have used the market data which we took at the Bloomberg terminal [http://www.bloomberg.com]. We have compared our results with the true last prices \( u_t (\tau) \) and \( u_t (2\tau) \). Let \( s_m = (s_b + s_a) / 2 \) be the mid point between bid and ask “today’s” prices of the underlying stock. First, we have calculated the minimizer \( u_{\alpha} (s, t) \) via the above minimization procedure, where \( (s, t) \in Q_{2\tau} \). This was done for every option for all days of its existence.

Those cases when \( u_{\alpha} (s_m, \tau) \) was between our extrapolated values of bid \( u_b (\tau) \) and ask \( u_a (\tau) \) prices, i.e. within the bid/ask spread, were not of any interest to us and we have not traded options in those days. However, we were only interested in those cases when our predicted price was larger than our extrapolated ask price for at least $0.02, i.e.
\( u_a (s_m, \tau) \geq u_a (\tau) + 0.03 \). Indeed, bid and ask prices are close to each other, which makes the case \( u_b (\tau) \leq u_a (s_m, \tau) \leq u_a (\tau) \) not interesting for trading. On the other hand, the cut-off value of $0.02 was chosen due to our computational experience. The same was for \( u_a (s_m, 2\tau) \), see our trading strategy below. In all cases of trading strategy listed below we buy and sell options just before the market closure.

**Trading Strategy:**

1. Suppose that \( u_a (s_m, \tau) \geq u_a (\tau) + 0.03 \) and also that \( u_a (s_m, 2\tau) \geq u_a (2\tau) + 0.03 \). Then we buy two (2) items of that stock option at \( t = 0 \), i.e. “today”. Next, we sell one item at \( t = \tau \) and sell the second item at \( t = 2\tau \). In doing so, we do not make any forecast at \( t = \tau \). Next, we apply our forecasting procedure at the day \( t = 2\tau \) as above and repeat the use of our trading strategy.

2. Suppose that \( u_a (s_m, \tau) \geq u_a (\tau) + 0.03 \) but \( u_a (s_m, 2\tau) < u_a (2\tau) + 0.03 \). Then we buy one (1) item of that stock option at \( t = 0 \). Next, we sell that item at \( t = \tau \). Next, we apply our forecasting procedure at the day \( t = \tau \) as above and repeat the use of our trading strategy.

3. Suppose that \( u_a (s_m, \tau) < u_a (\tau) + 0.03 \), but \( u_a (s_m, 2\tau) \geq u_a (2\tau) + 0.03 \). Then we buy one (1) item of that stock option at \( t = 0 \). Next, we sell that item at \( t = 2\tau \). Next, we apply our forecasting procedure at the day \( t = 2\tau \) as above and repeat the use of our trading strategy. However, we do not apply our forecasting procedure at the day \( t = \tau \).

4. Suppose that \( u_a (s_m, \tau) < u_a (\tau) + 0.03 \) and also that \( u_a (s_m, 2\tau) < u_a (2\tau) + 0.03 \). Then we neither buy nor sell this option “today”. Note that a particular case of this is the situation when both forecasted prices for \( t = \tau \) and \( t = 2\tau \) are within the bid/ask spread, i.e. when \( u_b (\tau) \leq u_a (s_m, \tau) \leq u_a (\tau) \) and also \( u_b (2\tau) \leq u_a (s_m, 2\tau) \leq u_a (2\tau) \).

We have evaluated twenty (20) liquid options, which were selected randomly among those options which are daily traded (see Introduction). Their short codes and our numbers for them are given in Table 2.

**Table 2. Short option codes.** “C” and “P” mean call and put options respectively. “Number of days” means the total number of days an option was evaluated by our procedure.
| Option Number | Option Short Code | Number of days |
|---------------|------------------|----------------|
| 1             | INTC US 01/17/15 C25 Equity | 181            |
| 2             | WFC US 11/22/14 C50 Equity | 52             |
| 3             | PFE US 11/22/14 C29 Equity | 68             |
| 4             | YHOO US 01/17/15 C50 Equity | 267            |
| 5             | AIG US 11/22/14 C55 Equity | 155            |
| 6             | AAPL US 01/17/15 C80 Equity | 130            |
| 7             | SBUX US 12/20/14 C80 Equity | 30             |
| 8             | AAL US 01/17/15 C40 Equity | 195            |
| 9             | QCOM US 12/20/14 C80 Equity | 31             |
| 10            | MSFT US 01/17/15 C45 Equity | 267            |
| 11            | QQQ US 11/22/14 C106 Equity | 45             |
| 12            | MRK US 11/22/14 C60 Equity | 58             |
| 13            | DDD US 01/17/15 C45 Equity | 119            |
| 14            | WMB US 02/20/15 C48 Equity | 43             |
| 15            | MSFT US 03/20/15 C50 Equity | 48             |
| 16            | SPY US 02/20/15 P205 Equity | 141            |
| 17            | IWM US 03/20/15 P110 Equity | 115            |
| 18            | EEM US 02/20/15 P39 Equity | 108            |
| 19            | YHOO US 03/20/15 C55 Equity | 76             |
| 20            | HYG US 03/20/15 P86 Equity | 48             |

Table 3 shows total profit/loss for each option resulting from our price forecast and the follow up application of our trading strategy. Losses are with the “−” sign. We did not invest any real money. Rather, for each option, we pretended that we buy and then sell the next day only one item and only at those days which were recommended by our trading strategy. To get results of this table, we have compared $u_a(s_m, \tau)$ with the true last price of the next day $u_l(\tau)$.

Table 3. Total gain/loss for each option of Table 2 during the evaluation period. The last row shows the total profit for all options.
| Option Number | Profit/loss |
|---------------|------------|
| 1             | 2.69       |
| 2             | 0.51       |
| 3             | 0.51       |
| 4             | 2.93       |
| 5             | 1.79       |
| 6             | 4.72       |
| 7             | 0.13       |
| 8             | 2.65       |
| 9             | 0.65       |
| 10            | 1.74       |
| 11            | 0.008      |
| 12            | 0.9        |
| 13            | −1.66      |
| 14            | −0.63      |
| 15            | 0.37       |
| 16            | 0.98       |
| 17            | −2.26      |
| 18            | 2.9        |
| 19            | 1.06       |
| 20            | 0.77       |
| **Total**     | **20.8**   |

The next interesting question is: *How accurate are we in our forecast?* Table 4 gives the average relative error of our forecast for each option. For each option, we have calculated the relative error for those days when we were selling that option by the above strategy. That relative error was

\[
\frac{\left| u_{\alpha}(s_m, \tau) - u_l(\tau) \right|}{u_l(\tau)}. \tag{4.1}
\]

Next, we took the average value of numbers (4.1) for each option of Table 2.

**Table 4.** The average relative accuracy of our forecast. The last row shows the total average relative accuracy over all options.
Table 4 shows that the average relative error (4.1) is rather large: 23.7%. We have observed that we often overestimated option prices. Nevertheless, our trading strategy enables us to be profitable at least on twenty options of Table 2.

## 5 Summary and conclusions

We have proposed a new mathematical model for the Black-Scholes equation. Instead of traditionally solving this equation backwards in time, which is a well posed problem, we solve it forwards in time, which is an ill-posed problem. The theory, which was previously developed in Klibanov (2014), guarantees the existence and uniqueness of regularized solution of this problem as the convergence of regularized solutions to the correct one as the level of the noise in the data $\delta$ and the regularization parameter $\alpha (\delta)$ tend to zero (Theorems 2, 3).

The best way to verify a mathematical model is to apply it to the real market data. And this is what we have done here. We have chosen randomly twenty (20) liquid options. Next, we have presented forecasts of option prices for the next two days for them. We have also developed a simple trading strategy, which is using our forecast to buy and sell options. We pretended that we buy and sell those options (the real money were not invested). The main conclusion, which can be drawn, is that even though our technique almost always overestimates option prices, still we got profits in seventeen (17) of above options and we got losses in three (3) of them. Furthermore, the total result is that, summing up these profits and losses, we are still profitable on our randomly selected twenty options.
We point out that we did not take into account the transaction cost. We conjecture that accounting for this cost would increase the above cut-off value of $0.02, while leaving the conclusion about the profit intact.

Even though our profits are small, this should not be discouraging. Indeed, we pretended to buy and then sell the next day only one-two items of an option at a time, and this was done only for twenty (20) options. However, large financial institutions buy and sell a large number of options daily. Hence, we conjecture that such large trading might indeed be profitable if using the technique of this paper. Still, we caution again that further studies of a much larger number of options should be performed to figure out how correct our conjecture is. However, we do not have enough resources to conduct such studies.

We believe that it is possible to refine our results and probably to increase potential profits. We are naturally concerned about the way of refining the accuracy of our predicted prices, see Table 4. One of such ways is to obtain more accurate values of the volatility coefficient, as compared with the implied volatility we have used here. Next, one should solve the problem (2.2)-(2.4) with this more accurate volatility coefficient. Most likely, the volatility depends on both stock price $s$ and time $t$, $\sigma = \sigma (s, t)$ . However, to compute the volatility, one needs to solve a very difficult coefficient inverse problem for the Black-Scholes equation. This problem is far more challenging than the one we consider here. Indeed, the solution $u = u (s, t, \sigma)$ of equation (2.2) depends nonlinearly on the coefficient $\sigma$. On the other hand, the problem, which was solved above, is linear.

The main challenge in solving nonlinear inverse problems is linked with the fact that the conventional least squares functionals for them are non convex, thus having many local minima and ravines, see, e.g. Isakov (2014) for this observation. Some ideas to calculate more accurate values of $\sigma (s, t)$ were proposed in Bouchouev and Isakov (1997), Bouchouev and Isakov (1999), Bouchouev, Isakov and Valdivia (2002) and Isakov (2014). However, these publications have not used our mathematical model. We believe that to work with our model, some modifications of currently known globally convergent numerical methods for coefficient inverse problems might be applied, see Klibanov (1997) (two papers), Klibanov and Thành (2014) and Klibanov and Kamburg (2015) for these methods.

Acknowledgment

The authors are grateful to Mr. Kirill Golubnichij for his help with some technical issues of creating the pdf file for this paper.

References

[1] Bouchouev, I. and Isakov, V. (1997): The inverse problem of option pricing, Inverse Problems, 13, L11-L17.

[2] Bouchouev, I. and Isakov, V. (1999): Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets, Inverse Problems, 15, R95-R116.

[3] Bouchouev, I., Isakov, V. and Valdivia, N. (2002): Recovery of volatility coefficient by linearization, Quantitative Finance, 2, 257-263.
[4] Hull, J.C. (2000): *Options, Futures, and Other Derivatives*, Prentice-Hall, New York.

[5] Isakov V. (2014): Recovery of time dependent volatility coefficient by linearization, *Evolution Equations and Control Theory*, 3, 1-16.

[6] Ivanov, V.K., Vasin, V.V. and Tanana, V.P. (2002): *Theory of Linear Ill-Posed Problems and Its Applications*, VSP, Utrecht.

[7] Klibanov, M.V. (1997): Global convexity in a three-dimensional inverse acoustic problem, *SIAM Journal on Mathematical Analysis*, 28, 1371-1388.

[8] Klibanov, M.V. (1997): Global convexity in diffusion tomography, *Nonlinear World*, 4, 247-265.

[9] Klibanov, M.V. and Th`anh, N.T. (2014): Recovering of dielectric constants of explosives via a globally strictly convex cost functional, Preprint, arXiv: 1408.0583 [math-ph], accepted for publication in *SIAM J. Applied Mathematics*.

[10] Klibanov, M.V. and Kamburg, V.G. (2015): Globally strictly convex cost functional for an inverse parabolic problem, Preprint, arXiv: 1502.01285 [math-ph].

[11] Klibanov, M.V. (2014): Carleman estimates for the regularization of ill-posed Cauchy problems, arxiv 1410.7521v1 [math-phys], accepted for publication in *Applied Numerical Mathematics*.

[12] Lavrentiev, M.M., Romanov, V.G. and Shishatskii, S.P. (1986): *Ill-Posed Problems of Mathematical Physics and Analysis*, AMS, Providence, RI.

[13] Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V. and Yagola, A.G. (1995): *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer, London.

[14] Wilmott P., Howison S. and Dewyne J. (1997): *The Mathematics of Financial Derivatives*, Cambridge University Press, New York.