# Powell Algorithm-Based Extraction of Backscatter Coefficients from LiDAR Signal Data

*Project - Optimization II*

Amelia Hoyos Vélez*[1]

[1]*Ingeniería Matemática, Universidad EAFIT, Medellín, Colombia*

November 12, 2024

## 1 Introduction

LiDAR, or Light Detection and Ranging, is a relatively obscure technology compared to sonar and radar, which have traditionally dominated the field of long-distance measurement tools. However, over the years, LiDAR has steadily expanded its influence across various fields such as atmospheric and oceanic sciences, autonomous vehicle development, mapping, precision agriculture, the prediction and evaluation of natural disasters, and astronomy [4]. As its applications grow, the need to develop accurate models that capture the behavior of LiDAR systems in different environments becomes increasingly important.

In this work, we present a method for constructing a theoretical model for a LiDAR signal based on an observed one. The goal is to estimate the backscatter coefficient $\beta(\lambda, z)$ of the medium such that the squared difference between the observed and modeled signals is minimized across the distance $z$.

To achieve this, we employ the Powell algorithm, a widely used approach in inverse problems for parameter estimation, particularly in optical applications [2]. The Powell algorithm is advantageous for this application because it does not rely on derivative calculations, which pose a significant challenge in the lidar models proposed due to their complexity and potential nonlinearity.

The remainder of this study is structured as follows. Section 2 provides an introduction to LiDAR technology. Section 3 describes the Powell algorithm. Section 4 presents the mathematical formulation of the optimization problem. Section 5, presents a sample problem and the results of applying the proposed method to construct a theoretical model. Finally, Section 6, discusses the findings, and Section 7 presents the conclusions.

## 2 LiDAR Technology

In simple terms, LiDAR (or lidar, as referred to in the remainder of this article) is a detection system similar to sonar and radar technologies, but it utilizes short-wavelength light waves from lasers instead of sound waves or radio waves [10]. While lidar generally offers higher precision than sonar or radar, it is typically more expensive.

*ahoyosv2@eafit.edu.co

The fundamental operation of a lidar sensor involves emitting beams of light toward a target or medium. The laser light interacts with the medium it encounters, causing a portion of the light to scatter or reflect back toward the sensor. The amount of backscattered light depends on the properties of the material the light traverses, as different materials scatter light in distinct ways. This reflected light is then captured by high-precision sensors. By measuring the time it takes for the light to return, the system can accurately calculate the distance to the target. Additionally, the intensity and pattern of the detected light can provide valuable information about the surface or medium the light has interacted with, including its texture, composition, and even molecular structure [3].

Several theoretical models describe the returned signal. The most common one is the *lidar equation* described in [5] [8], [7] which relates the received power $P(z)$ to the physical properties of the medium through which the light propagates. The lidar equation can be expressed as shown in (1):

$$P(z) = C\frac{P_0\beta(\lambda, z)}{z^2}\exp\left(-2\int_0^z \alpha(\lambda, z')\,dz'\right), \tag{1}$$

where:

- $P(z)$ is the received signal power from a given distance $z$,
- $P_0$ is the initial power of the emitted laser,
- $C$ is a calibration constant of the laser,
- $\beta(\lambda, z)$ is the backscatter coefficient at wavelength $\lambda$ and distance $z$,
- $\alpha(\lambda, z)$ is the absorption coefficient at wavelength $\lambda$ and distance $z$.

The lidar equation uses two important coefficients to describe how the medium interacts with the laser light: the backscatter coefficient $\beta(\lambda, z)$ which indicates the fraction of light scattered back when interacting with the material and the absorption coefficient $\alpha(\lambda, z)$ which indicates the fraction of light absorbed by the material as it travels through it.

Although the lidar equation provides a relatively simple and accurate model, obtaining precise values for the backscatter and absorption coefficients in real-world applications presents significant challenges. In atmospheric studies, for example, these coefficients vary widely depending on factors such as location, weather conditions, and time, making it difficult to assign exact values.

For this reason, optical research typically focuses on analyzing real signal data rather than attempting to fully model the instrument itself. Nevertheless, developing a mathematical model for lidar behavior is crucial for advancing knowledge in this field.

This paper presents a method that, given real signal data, allows the estimation of the backscatter coefficient. Then to compute the extinction coefficient one can use the Klett Inversion Method that states that there exists a ratio $S(z)$ called lidar ratio such that (2) holds.

$$\alpha(z) = S(z) \cdot \beta(z) \tag{2}$$

By determining both the backscatter and extinction coefficients, one can fully model a lidar signal. These methods will streamline lidar signal modeling, enabling deeper investigations into lidar systems and advancing the field of instrumentation science.

# 3 The Powell Algorithm

To solve an inverse problem, where the goal is to estimate parameters within a model, one typically begins by assigning initial values to the parameters and then iteratively adjusts them to minimize the difference between the model and the observed data. The enhanced Powell algorithm follows this principle. Named after M. J. Powell [9], this algorithm finds the minimum of an objective function without relying on gradient computations, instead using an improved direction-set method with sophisticated line minimization.

Let $N$ denote the maximum number of iterations, $tol$ the position tolerance, and $ftol$ the function value tolerance. Let $x^{(0)}$ represent the initial guess for the minimum of the continuous cost function $f(x)$. The algorithm begins by determining the number $n$ of linearly independent directions, initially chosen as the unit vectors as shown in (3).

$$\mathbf{d} = [\mathbf{d_1}^T, ..., \mathbf{d_n}^T] = I_n. \tag{3}$$

At each iteration $k = 1, ..., N$, the algorithm performs a line search to minimize the function along each direction $d_i$ starting on $x^{(k)}$, yielding an optimal step size vector $\alpha^{(k)}$. The line search process consists of two main phases: bracketing and refinement using the golden section method. Bracketing is the process of finding an interval $[\alpha_l^{(k)}, \alpha_h^{(k)}]$ that contains the optimal step size $\alpha^{(k)}$ of the function along the search direction. Let $\alpha_0^{(k)}$ be an initial starting point for $\alpha^{(k)}$. We define $\alpha_l = 0$ and $\alpha_h = \alpha_0^{(k)}$.

The algorithm expands this interval by doubling $\alpha_h$ until it finds three points where the middle point has a lower function value than both endpoints. This creates a bracket as the function must have a minimum within this interval.

Once we have bracketed the minimum, we use the golden section method to find its location. This method is named after the golden ratio $\phi$ shown in (4).

$$\phi = \frac{1 + \sqrt{5}}{2} \approx 1.618, \quad \gamma = 2 - \phi. \tag{4}$$

At each iteration, we evaluate two interior points as shown in (5).

$$\begin{aligned} \alpha_1 &= \alpha_l + \gamma(\alpha_h - \alpha_l), \\ \alpha_2 &= \alpha_h - \gamma(\alpha_h - \alpha_l). \end{aligned} \tag{5}$$

These points divide the interval into sections with ratio $\phi$. We then evaluate the function at $f(x^{(k)} + \alpha_1 \mathbf{d})$ and $f(x^{(k)} + \alpha_2 \mathbf{d})$. Depending on their values, we update $\alpha_l$ and $\alpha_h$ to reduce the interval where the minimum is located. It is important to note that $\alpha_1$ and $\alpha_2$ have a unique property: when we remove one section, the remaining interval can be divided in the same proportion while reusing one of the previous evaluation points. This minimizes the number of function evaluations needed. The process continues until the interval is sufficiently small, when $|\alpha_h - \alpha_l| < tol$. The optimal step size vector over all directions $\vec{\alpha}^{(k)}$ is then used to compute the new point as shown in (6).

$$x^{(k)} = x^{(k-1)} + \vec{\alpha}^{(k)} \mathbf{d}. \tag{6}$$

To prioritize directions based on their contribution to the objective function's improvement, we calculate the improvement for each direction $i$ as shown in (7). Directions are then sorted in an array $A$ from most improvement to least improvement.

$$\text{improvement}_i = f(x^{(k-1)}) - f(x_i^{(k)}). \tag{7}$$

The direction that leads to the least improvement is replaced by a new direction $\mathbf{d_n}$, defined as the normalized difference between the current and previous states as shown in (8).

$$\mathbf{d_n} = \frac{x^{(k)} - x^{(k-1)}}{\|x^{(k)} - x^{(k-1)}\|}. \tag{8}$$

This new direction aims to incorporate recent changes and guide the optimization process in an updated trajectory. It ensures that the search space evolves effectively to maintain or improve the convergence properties of the algorithm.

We then apply the Gram-Schmidt process to the directions in $A$ to ensure that the directions $\mathbf{d_i}$ are mutually orthogonal. The orthogonalization proceeds as shown in (9).

$$\begin{aligned} &\text{for } i = 1 \text{ to } n : \\ &\text{for } j = 1 \text{ to } i-1 : \\ &\quad \mathbf{d_i} = \mathbf{d_i} - (\mathbf{d_i} \cdot \mathbf{d_j})\mathbf{d_j}, \\ &\quad \mathbf{d_i} = \frac{\mathbf{d_i}}{\|\mathbf{d_i}\|} \quad \text{if } \|\mathbf{d_i}\| > 10^{-10}, \end{aligned} \tag{9}$$

where $\|\mathbf{d_i}\|$ represents the norm of direction $\mathbf{d_i}$. The condition $\|\mathbf{d_i}\| > 10^{-10}$ is included to prevent division by near-zero values, ensuring numerical stability.

Throughout the optimization process, the algorithm maintains the best solution found so far (10).

$$x* = \arg\min_{i \leq k} f(x^{(i)}). \tag{10}$$

The algorithm takes into account two convergence criterion, shown in (11) or (12).

$$|f^{(k)} - f^{(k-2)}| < ftol. \tag{11}$$

$$\|x^{(k)} - x^{(k-1)}\| < tol. \tag{12}$$

The algorithm terminates when any convergence criterion is met or when the maximum number of iterations is reached.

# 4 Problem Formulation

In this study, we assume the use of a monochromatic laser beam, implying that the backscatter and absorption coefficients are independent of wavelength. Let the signal data $P_{\text{measured}}(z)$ be known over a range $[0, z_f]$. Let $P(z)$ be a model for our lidar signal on that same range based on (1), where $P_0$ and a calibration constant $C$ are known. The objective function is defined in (13) as the integral of the square of the difference between these two functions.

$$J = \int_0^{z_f} \left(P(z) - P_{\text{measured}}(z)\right)^2 \, dz. \tag{13}$$

Note that, as shown in (1), $P(z)$ depends on the backscattering and absorption coefficients. We initialize $\alpha(z) = 0$ and optimize the cost function with respect to $\beta(z)$. Once the backscatter coefficient is determined, one can construct a theoretical model for the lidar signal.

More formally, the optimization problem is defined as show in (14).

$$\begin{aligned} &\text{minimize } J = \int_0^{z_f} \left(P(z) - P_{\text{measured}}(z)\right)^2 \, dz \\ &\text{subject to:} \\ &P(z) = \frac{P_0 C \beta(z)}{z^2} \exp\left(-2 \int_0^z \alpha(z') \, dz'\right) \\ &\beta(z) \geq 0. \end{aligned} \tag{14}$$

To approximate the integral in the objective function, we apply Simpson's rule. For this purpose, we divide the distance interval $[0, z_f]$ into $N$ subintervals, where $z_i = i \cdot \Delta z$ for $i = 0, 1, \ldots, N$, and $\Delta z$ is the uniform step size. The integral $J(z)$ can then be approximated as shown in (15).

$$J \approx \frac{\Delta z}{3} \left[ f(0) + 4 \sum_{i=1,\text{ odd}}^{n-1} f(z_i) + 2 \sum_{i=2,\text{ even}}^{n-2} f(z_i) + f(z_n) \right], \tag{15}$$

where $f(z) = \left(P(z) - P_{\text{measured}}(z)\right)^2$ is evaluated in the discretized points.

To handle the constraint $\beta(z) \geq 0$, we apply a penalty method. This introduces a penalty term $\phi(z)$ to the objective function, defined in (16).

$$\phi = \sum_{i=0}^{n} \max(0, -\beta(z_i))^2. \tag{16}$$

The modified optimization problem is then shown in (17).

$$\text{Minimize } J_{\text{penalized}} = J + \rho \cdot \phi, \tag{17}$$

where $\rho$ is a large positive constant to enforce the constraints.

To solve this optimization problem we used the Powell Algorithm.

# 5 Sample Problem and Results

The data used to demonstrate the utility of this method is from the High spectral resolution lidar [6], [1] from the day Thu, Sep 26, 2019. This lidar scans clouds and aerosols in the atmosphere with a vertical resolution of 15m. The instrument operates at multiple wavelengths, including 532nm, 355nm, and 1064nm, each providing specific insights into particle size, concentration, and type. For this study, we focus on the 532nm wavelength.

The file used has different data slices, we will use data slice 100 as our case study. The original signal and predicted signal of our model is shown in 1. The original and the extracted backscattering coefficient is shown in 2.

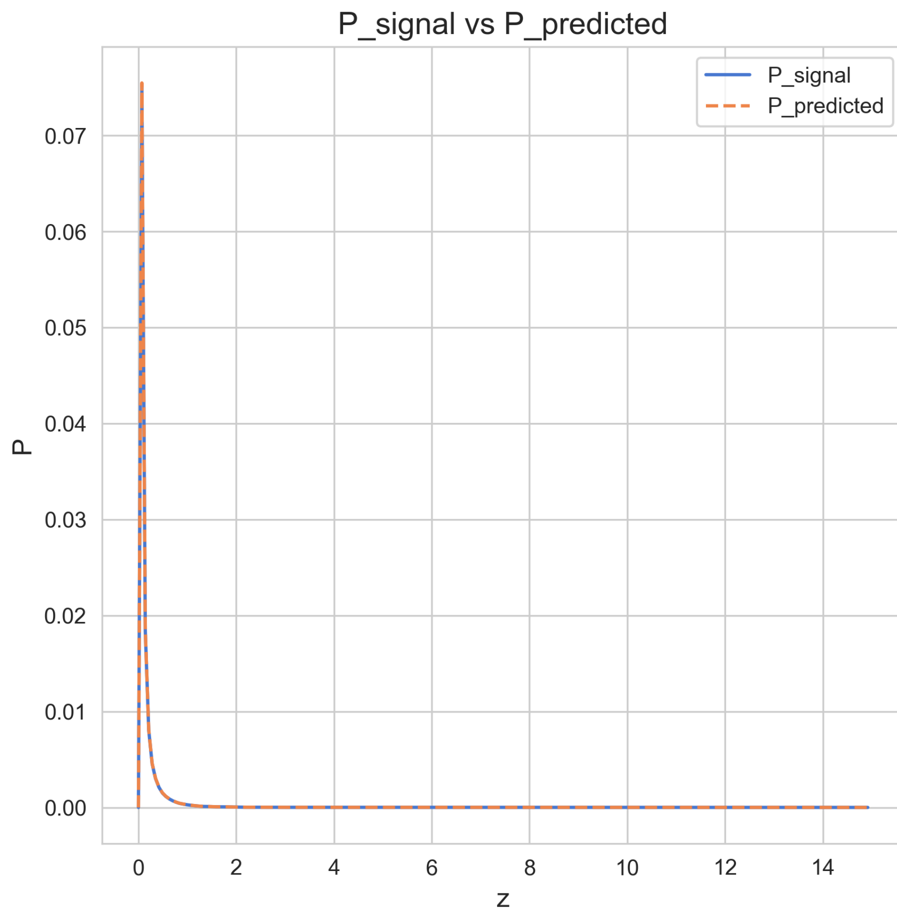


Figure 1: original and Predicted Lidar Signal for Data Slice 100

The total cost $J$ of this solution, which quantifies the overall discrepancy between the observed and predicted lidar signals, is calculated to be $J = 4.849 \times 10^{-9}$. This value reflects the minimized error in fitting our model to the actual lidar data, with a very low cost indicating a strong agreement between the model and the observed data.

Furthermore, the squared difference between the actual backscattering coefficient and the one estimated by our model is $1.430 \times 10^{-7}$. This small difference suggests that our model accurately approximates the lidar backscattering profile, capturing essential atmospheric features and variations in the data with a high degree of precision.

Overall, these metrics indicate that the model effectively reproduces the behavior of the real lidar signal by accurately estimating the backscattering coefficient. The method is both robust and

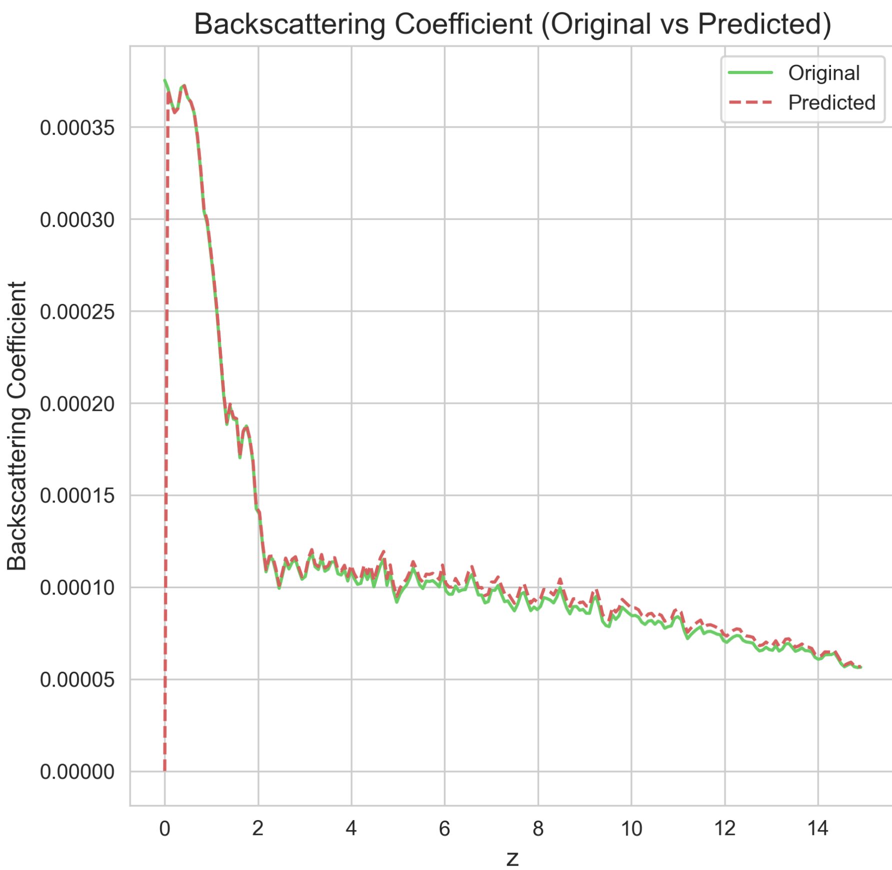


Figure 2: Original and Predicted Lidar Backscattering coefficient for Data Slice 100

well-suited for interpreting lidar measurements, providing reliable parameter estimates.

We extended the backscattering coefficient estimation to an additional 94 data slices. With 95% confidence, we determined that the cost function $J$ falls within the interval $(-1.928 \times 10^{-6}, 3.186 \times 10^{-4})$, indicating a consistently small cost and thus a close fit between the modeled and actual lidar signals.

Similarly, the squared difference between the actual and predicted backscattering coefficients lies within the range $(-1.391 \times 10^{-7}, 5.059 \times 10^{-6})$. This narrow confidence interval demonstrates the model's accuracy in estimating backscattering across diverse data slices, confirming the model's robustness and adaptability to variations in lidar signal characteristics. We conclude that it generalizes effectively across different atmospheric conditions and lidar measurements.

# 6 Discussion

The results demonstrate the effectiveness of the Powell algorithm in extracting backscatter coefficients from LiDAR signal data. The narrow confidence interval of the cost function across 94 data slices indicate strong consistency in the method's ability to minimize the cost function. Moreover, the comparison between predicted and actual backscattering coefficients, with a squared difference of suggests high accuracy in the parameter estimation.

This approach effectively addresses the inherent nonlinearity of the lidar equation, achieving high accuracy without the need for derivative-based calculations. This accuracy is especially impressive considering the usual complexities associated with atmospheric lidar data, where signal variability and noise often present significant challenges. The method's performance with the NASA ORACLES HSRL2 dataset highlights its potential for real-world lidar applications, particularly in atmospheric research. Further exploration with diverse datasets would be valuable to fully validate its robustness and applicability across different environmental conditions.

Future work could focus on extending this method to also calculate the extinction coefficient without depending on the Klett Inversion Method. Additionally, the approach could be adapted for real-time processing of LiDAR data in applications such as autonomous navigation and atmospheric monitoring.

# 7 Conclusions

This study presents a novel approach to extracting backscatter coefficients from LiDAR signal data using the Powell algorithm. These results suggest that the Powell algorithm-based approach provides a reliable method for constructing theoretical LiDAR signal models from experimental data, particularly for extraction the backscattering coefficient of the lidar system.

# 8 Implementation

The implementation of this algorithm can be found at this GitHub repository.

# References


[1] S. P. Burton et al. “Calibration of a high spectral resolution lidar using a Michelson interferometer, with data examples from ORACLES”. In: *Appl. Opt.* 57.21 (July 2018), pp. 6061–6075. DOI: 10.1364/AO.57.006061. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-57-21-6061.

[2] Zhonghao Chang et al. “Application of the Powell algorithm for estimating optical properties of a semitransparent medium based on time-domain information”. In: *Appl. Opt.* 62.36 (Dec. 2023), pp. 9493–9501. DOI: 10.1364/AO.504903. URL: https://opg.optica.org/ao/abstract.cfm?URI=ao-62-36-9493.

[3] T. Gasmi. *An Introduction to Lidar and Its Applications*. International Centre for Theoretical Physics, 2008. URL: https://indico.ictp.it/event/a08142/session/48/contribution/31/material/0/0.pdf.

[4] GISGeography. *13 Lidar Uses and Applications*. 2022. URL: https://gisgeography.com/lidar-uses-applications/.

[5] Duane Hardesty. *The Lidar Equation*. Accessed: 2024-11-10. 1997. URL: https://lidar.ssec.wisc.edu/papers/dhd_thes/node10.htm.

[6] Chris Hostetler and Rich Ferrare. *ORACLES HSRL-2 Data Archive*. Database. Data collected during the ORACLES mission (8/19/2016-10/27/2018) using the NASA P-3 Orion aircraft. NASA Earth Science Project Office, 2023. DOI: 10.1364/AO.47.006734. URL: https://espoarchive.nasa.gov/archive/browse/oracles/P3/HSRL2.

[7] *Lidar Remote Sensing Overview*. 2007. URL: https://superlidar.colorado.edu/Classes/Lidar2007/Lecture04.pdf.

[8] Curtis D. Mobley. *The Lidar Equation*. Accessed: 2024-11-10. 2021. URL: https://oceanopticsbook.info/view/radiative-transfer-theory/level-2/the-lidar-equation.

[9] M. J. Powell. “A new algorithm for unconstrained optimization”. In: *Nonlinear Programming*. Elsevier, 1970, pp. 31–65.

[10] *What is Lidar and What is it Used For?* National Oceanic and Atmospheric Administration, n.d. URL: https://oceanservice.noaa.gov/facts/lidar.html.