Optimization of an One Dimensional Hyper telescope for a Direct Imaging in Astronomy.

Abstract :

We describe an application of nonlinear optimization in interferometric optical astronomy. The aim is to find the relative positions of the output pupils and the modulus of the beams through each pupil of a linear array of telescopes in order to design an instrument capable of imaging exoplanets. The problem is modelized under the form of a semi-infinite nonlinear minimization problem. The model problem is transformed by a simple discretization into a minimization problem with a finite number of constraints, then it is solved by using a minimization solver. Numerical experiments are reported.

What is Non Linear Optimization ?

Non-linear optimization refers to the process of finding the optimal solution (minimum or maximum) of an objective function that depends on multiple variables, where the relationship between the variables and the objective function is non-linear. Non-linear optimization problems can be challenging to solve because they may have multiple local minima or maxima, making it difficult to find the global optimum.

To perform non-linear optimization, we use different algorithms and techniques, such as:

- Gradient Descent
- Newton's Method
- Genetic Algorithms
- Simulated Annealing

What is Hyper telescope and how it is used ?

A hyper-telescope is an advanced type of astronomical interferometer that combines the light from a large number of smaller telescopes or mirrors to achieve extremely high angular resolution and sensitivity.
The main idea behind a hyper-telescope is to distribute many small mirrors or telescopes over a vast area, effectively creating a giant, virtual telescope with an aperture equivalent to the maximum distance between the individual elements. This large effective aperture allows the hyper-telescope to achieve much higher angular

resolution than conventional telescopes, enabling it to resolve details in distant celestial objects that would otherwise be impossible to observe.
In addition to its high angular resolution, a hyper-telescope has the potential to achieve high sensitivity, as the combined light-gathering power of all its individual elements is greater than that of a single, larger telescope. This increased sensitivity would allow the hyper-telescope to detect fainter objects and study their properties in greater detail.
One of the main challenges in building a hyper-telescope is the need for precise control and alignment of the individual elements, as well as the accurate combination of the light from all elements to form a single, coherent image. Various designs and techniques have been proposed to overcome these challenges, and some small-scale prototypes have been built to demonstrate the feasibility of the hyper-telescope concept.
Potential applications of hyper-telescopes include imaging the surfaces of stars, studying the formation and evolution of galaxies, and directly observing.

What is Interferometric Optical Astronomy and why we use it ?

Interferometric optical astronomy is a technique used in observational astronomy that combines the light collected from two or more separate telescopes to achieve greater angular resolution than what would be possible with a single telescope. The technique is based on the principle of interferometry, which involves the superposition of light waves to extract information about the source of the light.

In interferometric optical astronomy, the light from a celestial object is collected by two or more telescopes, often separated by a significant distance. The light collected by each telescope is combined to form an interference pattern, which contains information about the spatial distribution, structure, and other properties of the celestial object.
The reason for using interferometry in optical astronomy is to overcome the diffraction limit, which is the fundamental resolution limit imposed by the wave nature of light. The angular resolution of a telescope is limited by the ratio of the wavelength of the observed light to the diameter of the telescope's aperture. By combining the light from two or more telescopes, interferometric optical astronomy effectively increases the aperture size, resulting in higher angular resolution and the ability to observe finer details in celestial objects.

One well-known example of an optical interferometer is the Very Large Telescope Interferometer (VLTI) in Chile, which combines the light from four 8.2-meter telescopes and several auxiliary telescopes to achieve a resolution equivalent to a single telescope with a diameter of up to 130 meters.
Interferometric optical astronomy has been instrumental in studying various astronomical objects and phenomena, such as the shapes and sizes of stars, the atmospheres of exoplanets, the structure of active galactic nuclei, and the formation and evolution.

What is the Aim of this Paper ?

Our Aim here is to optimize the design of a one-dimensional hypertelescope, which is an array of individual telescopes that work together to improve the resolution and dynamic range of the resulting image. Two important factors they consider in this optimization process are the "optimal modulus (amplitude) of each light beam" and the "relative positions of the output pupils."

- **Optimal modulus (amplitude) of each light beam :** Each telescope in the array collects light from the observed astronomical object. The light waves from each telescope interfere with each other when combined, and the amplitude of each light wave (also referred to as the modulus) is a crucial factor in this interference pattern. By optimizing the amplitude of each light beam, the authors aim to maximize the constructive interference and minimize the destructive interference, which can lead to clearer, higher-resolution images.

- **Relative positions of the output pupils :** The output pupils are the points in the optical system where the light collected by the individual telescopes is combined to create the final image. The relative positions of these output pupils determine how effectively the light from the individual telescopes can be combined, which in turn influences the resolution and dynamic range of the resulting image. By optimizing the relative positions of the output pupils, the authors aim to achieve the best possible image quality from the hypertelescope.

In summary, here we are trying to find the best combination of light beam amplitudes and output pupil positions in a one-dimensional hypertelescope to achieve high-resolution and high dynamic range images, which is essential for observing distant and faint astronomical objects, such as exoplanets.

How to Solve this ?

Here we follow a three-step approach to address the optimization problem:

1. Model the problem as a semi-infinite nonlinear minimization problem, which represents the initial complex and challenging nature of the problem.

A semi-infinite nonlinear minimization problem is an optimization problem characterized by the following features:

- Nonlinear objective function and/or constraints: In the problem, either the objective function or the constraints, or both, are nonlinear. This means they cannot be represented by linear equations or inequalities. Nonlinear problems are generally more challenging to solve than linear ones, as they may have multiple local minima, making it harder to find the global minimum.

- Semi-infinite constraints: The problem has an infinite number of constraints. These constraints often involve continuous functions, integrals, or other mathematical expressions with an infinite domain. In contrast, a problem with a finite number of constraints has a limited number of equations or inequalities that define the feasible region.

Solving a semi-infinite nonlinear minimization problem directly can be computationally challenging due to its complexity. In practice, researchers often apply techniques like discretization or approximation methods to transform the problem into a more manageable form with a finite number of constraints before attempting to solve it. By transforming the problem into a more manageable form, it becomes easier to apply optimization algorithms or solvers to find an optimal solution.

2. Transform the problem into a more manageable form by applying discretization, resulting in a minimization problem with a finite number of constraints. Discretization is a process of converting a continuous problem into a discrete (finite) one by breaking it into smaller, finite components. In this case, the authors convert the semi-infinite problem into a minimization problem with a finite number of constraints, which is easier to handle computationally.

3. Use a minimization solver (an optimization algorithm) to find the optimal solution, which provides the desired moduli (amplitudes) and positions of the output pupils necessary for high-resolution and high dynamic range imaging in the hypertelescope.

By following this approach, the authors aim to design a one-dimensional hypertelescope that can directly image exoplanets and other astronomical objects with high resolution and dynamic range.

Motivation for this Paper in Brief.

Traditional interferometric techniques in optical astronomy rely on indirect methods for image restoration, requiring numerical reconstruction of the image from partial information. This approach does not allow for the selection of light from a single pixel for direct spectral analysis. The hypertelescope concepts proposed by various researchers address this issue by enabling direct imaging through specific conditioning of the light beams coming from an array of telescopes.

This optimization of the design of a one-dimensional hypertelescope aims to achieve high resolution and dynamic range in the image, which are essential for observing faint and distant astronomical objects, such as exoplanets orbiting distant stars. By developing an optimized hypertelescope, the authors hope to contribute to advancements in astronomical imaging, enabling better observation and analysis of exoplanets and other celestial objects.

The aforementioned information provides an overview of the paper that we are currently handling.

Introduction :

Now lets discuss about the detailed Structure of a Hyper-Telescope, as discussed before a hyper-telescope is an advanced type of astronomical interferometer that combines the light from a large number of smaller telescopes or mirrors to achieve extremely high angular resolution and sensitivity. A hyper-telescope consists of several key components and has a structure designed , Here's an overview of the main elements of a hypertelescope :

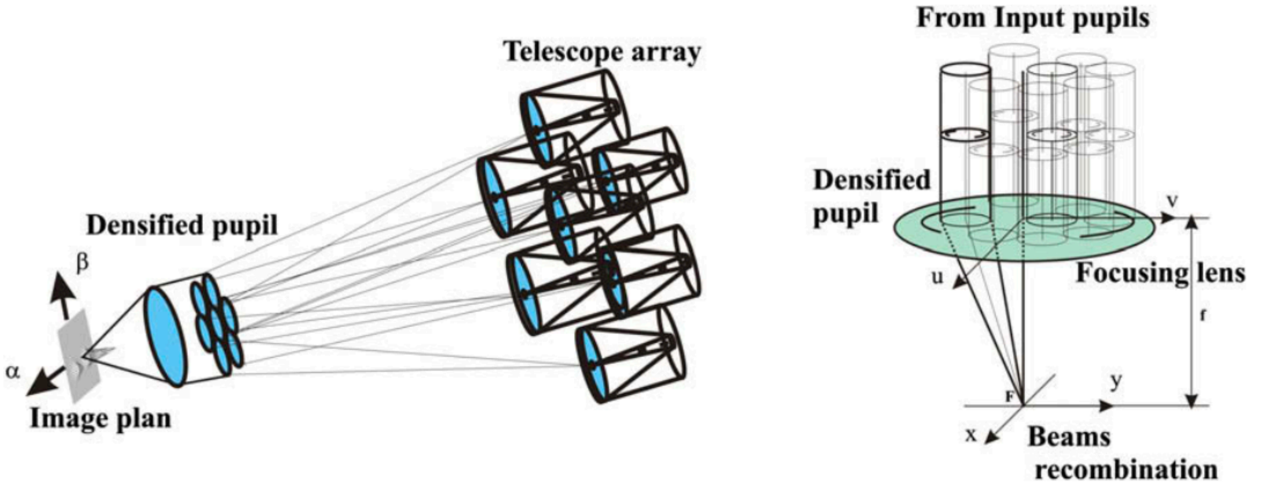

Fig. 1. Structure of a hypertelescope.

I. **Array of mirrors or telescopes :** The primary component of a hypertelescope is a large array of smaller mirrors or telescopes, distributed over a vast area. Each individual mirror or telescope collects light from a celestial object and redirects it towards a central focal point or a beam combiner. The large separation between the individual elements creates an effective aperture that is much larger than any single telescope, resulting in high angular resolution.

II. **Beam combiner :** The beam combiner is a crucial part of a hypertelescope, responsible for combining the light collected by the individual mirrors or telescopes. This is typically done using a system of delay lines, beam splitters, and additional optics to ensure that the light waves from each element are aligned and in phase when they interfere with each other. The interference pattern generated by the beam combiner contains information about the spatial structure of the observed object.

III. **Pupil densifier :** A pupil densifier is an optional component used in some hypertelescope designs to concentrate the light from the individual mirrors or telescopes into a smaller area. This can increase the sensitivity of the hypertelescope and reduce the size of the beam combiner and other downstream components.

IV. **Imaging camera or detector :** The final component of a hypertelescope is an imaging camera or detector that captures the interference pattern generated by the beam combiner. This data is then processed using specialized algorithms to reconstruct an image of the observed object with high angular resolution.

V. **Control and alignment systems:** To maintain the precise alignment and positioning of the individual mirrors or telescopes, a hypertelescope employs advanced control and alignment systems. These systems ensure that each element is accurately pointed towards the celestial object being observed and that the optical path lengths between the individual elements and the beam combiner are precisely maintained. The control and alignment systems typically involve a combination of active and passive mechanisms, including actuators, sensors, and feedback loops.

VI. **Data processing and image reconstruction :** The raw data captured by the imaging camera or detector needs to be processed to produce a final, high-resolution image of the observed object. This involves the use of specialized algorithms and software to analyze the interference pattern, remove any noise, and reconstruct the image. This process can be computationally intensive and may require significant computing power to handle the large datasets generated by a hypertelescope.

In summary, a hypertelescope has a complex structure consisting of an array of mirrors or telescopes, a beam combiner, a pupil densifier, an imaging camera or detector, control and alignment systems, and data processing and image reconstruction components. These elements work together to provide extremely high angular resolution and sensitivity, enabling the hypertelescope to observe distant celestial objects with unprecedented detail.

The cophasing in hyper-telescope refers to the process of aligning the light waves collected by the individual mirrors or telescopes so that they are in phase when combined at the beam combiner. When the light waves are in phase, they interfere constructively, resulting in a strong and coherent signal that contains information about the observed object and it has to be achieved with a sub-micrometric accuracy.

These processes are taken carefully and in resulting to achieve high resolution of the images because some times the angular separation between a star and an exoplanet is expected to be in range of nano radian and the ratio of intensities can be less than 10^6,this results can be reached in a limited field and number of pixels remains low.

In this preliminary study, we will consider a linear array of tele- scopes corresponding to pupils arranged along a straight line. Applied optimization to the design of an optical instrument capable to image exoplanets has been already investigated by Kasdin et al. [3] and Vanderbei [10]. As in Vanderbei approach, we propose to formulate the design problem as a semi-infinite minimization problem. After discretization, the model problem is transformed into a con- strained nonlinear minimization problem. Then, the problem is solved by using a nonlinear optimization solver.

Densified Pupil and Point Spread Function (PSF) :

In this section we will look more deeply into what is densified pupil and point spread function and how these to are related to each other and in our optimization problem. As we already discussed above what a pupil densifier does lets what is it mean to say the corresponding Point Spread Function. The point spread function describes the response of an imaging system to a point source, such as a star. In the context of a hypertelescope or an astronomical interferometer, the PSF represents the shape and intensity distribution of the image produced by the instrument when observing a point source. The PSF is a crucial concept in understanding the quality of the images produced by the instrument, as it provides information about the resolution, contrast, and other imaging properties.now lets learn the relationship between the densified pupil and the point spread function.

The Optical Field of a Monochromatic plane wave through a pupil is characterized by a modulus(amplitude) and a phase (phase represents the position of the wave in its oscillation cycle),we assume here all the beams have the same phase (cophasing).The optical field of n pupils, centered at (u_k, v_k) for k = 1, ..., n, and with the same diameter d, is given by the function $g(u, v)$. In this function, a_k is the modulus (amplitude) of beam k, and 1_{B_k} is the characteristic function of the closed disk with centre (u_k, v_k) and diameter d. The characteristic function, 1_{B_k} , is equal to 1 inside the disk and 0 outside the disk.

The optical field $g(u, v)$ is defined as the sum of the individual beams from each of the n pupils. The function $g(u, v)$ represents the combined optical field of all the pupils, which is essential for understanding the imaging properties of the system and analyzing the point spread function (PSF).

$$g(u, v) = \Sigma_{k=1}^{n} a_k \ 1_{B_k}(u, v)$$

In the image field, the optical field is the Fourier Transform of the function g and is defined by

$$\hat{g}(x, y) = \int\int g(u, v) \ e^{-\frac{2i\pi}{\lambda f}(xu+yv)} du \ dv$$

Where λ is the wavelength and f is the focal length.The image plane intensity is given by the point spread function and is defined to be the square of modulus of $\hat{g}$.Therfore the PSF is given by :

$$\Psi(x, y) = \frac{|\hat{g}(x, y)|^2}{|\hat{g}(0,0)|^2}$$

In the below graph of a normalised PSF obtained with four pupils. The aim is to find the moduli a_k and positions (u_k, v_k) of the pupils such that the main central lobe of the PSF is as narrow as possible for high resolution, and the secondary lobes are as low as possible to achieve a strong dynamic range (the ability to distinguish between bright and faint objects in the image).

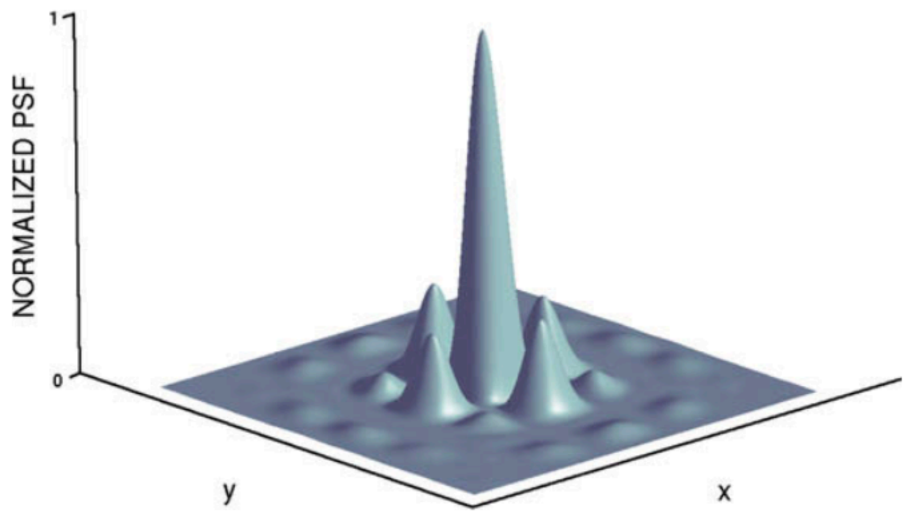

Fig. 2. Normalized PSF with four pupils.

Let B be a closed disk with the centre at position 0 and diameter d. s represents the surface of one pupil with diameter d. Since all pupils have the same diameter, the normalized PSF can be written as :

$$\Psi(x,y) = \frac{1}{s}\ |\hat{h}(x,y)|^2\ \frac{|\Sigma_{k=1}^{n}\ a_k\ e^{-\frac{2i\pi}{\lambda f}(xu_k+yv_k)}|^2}{|\Sigma_{k=1}^{n} a_k|^2}$$

This equation represents the combined PSF of all pupils, with each term in the sum corresponding to the contribution of one pupil. The sum takes into account the amplitude a_k and the position (u_k, v_k) of each pupil.

$\hat{h}(x,y)$ is the PSF associated with only one pupil, which is calculated using the double integral expression given below . This integral represents the Fourier transform of the characteristic function of the closed disk B over the spatial frequencies (x, y).

$$\hat{h}(x,y) = \int\int e^{-\frac{2i\pi}{\lambda f}(xu+yv)}\ 1_B(u,v)\ du\ dv$$

Since the normalized PSF is homogeneous with respect to a_k , we can assume the sum of all a_k values is equal to 1 ,This normalization constraint simplifies the optimization problem.

$$\Sigma_{k=1}^{n} a_k = 1$$

Our Assumption that the sum of all a_k values equal to 1, simplifies the optimization problem by reducing the number of free parameters that need to be considered when searching for the optimal configuration of the pupils.

Without the normalization constraint, you would have to optimize the moduli (a_k) and positions (u_k, v_k) for each pupil independently, resulting in a larger search space and potentially more complex optimization problem. By constraining the sum of a_k values to equal 1, you effectively reduce the degrees of freedom in the problem, as the value of one modulus (e.g., a_n) can be determined from the values of the others $(a_1, a_2, , , , , , a_{(n-1)})$. This reduces the size of the search space and makes the optimization problem easier to solve.

In addition to simplifying the optimization problem, the normalization constraint also ensures that the total power (or intensity) of the PSF remains constant. This is important when comparing different configurations of the pupils, as it allows you to focus on optimizing the shape and properties of the PSF (such as the width of the central lobe and the height of the secondary lobes) without having to worry about the overall intensity of the PSF.

Linear Array of Telescope :

In this paper we consider the particular case of linear array of telescopes, that is $v_k = 0$ for all $k = 1,2,3\ldots,n$, considering the above assumption the Normalized PSF becomes

$$\Psi(x, y) = \frac{1}{s}\ |\hat{h}(x, y)|^2\ |\Sigma_{k=1}^{n}\ a_k\ e^{-\frac{2i\pi}{\lambda f}(xu_k)}|^2$$

We can see that in the above equation the Optimization parameters a_k and u_k doesn't depend on y. So we set y to be some arbitrary value. Let us set $y = 0$. By introducing the polar coordinates $u = r\ cos\theta$ and $v = r\ sin\theta$, the value of $\hat{h}$ at $(x,0)$ can be written as

$$\hat{h}(x,0) = \int_0^{\frac{d}{2}} \int_0^{2\pi} e^{-\frac{2i\pi}{\lambda f}(x\ r\ cos\theta)}\ r\ d\theta\ dr$$

Now by using change of Variables $\alpha = \dfrac{d}{\lambda f}x$, we can write $\hat{h}$ as a function of α , that is

$$\hat{h}(\alpha) = \frac{2s}{\pi\alpha}J_1(\pi\alpha)$$

Where J_1 is the first order Bessel function of the first kind, now we can write the Normalized PSF as the function of α.

$$\Psi(\alpha) = |\frac{2}{\pi\alpha} J_1(\pi\alpha)|^2 \, |\Sigma_{k=1}^{n} a_k \, e^{-\frac{2i\pi u_k}{d}\alpha}|^2$$

Here we simplified the expression for the normalized PSF by considering a linear array of telescopes and setting y to an arbitrary value (in this case, 0). This simplification, along with the introduction of polar coordinates and the change of variables, leads to a more manageable expression for $\Psi(\alpha)$. This simplified expression can be used to analyze and optimize the configuration of the telescopes in the linear array, ultimately aiming to improve the imaging performance of the system.

Optimization Model :

Here we formulate the problem as follows. We must find the moduli a_k and the positions u_k for which the PSF is as small as possible in a region close to the main central lobe. We define an interval $[\alpha_{min} \, , \, \alpha_{max}]$ of α values for which we want that the values $\Psi(\alpha)$ are as small as possible. It is in this interval that the main central lobe of a secondary source of light could be detected. We call it the clean field of view (CLF).

Minimizing the PSF does not directly result in a narrow central lobe. However, the objective of the optimization problem discussed in previous passages is to find a configuration of telescopes that results in a narrow central lobe and low secondary lobes. The optimization problem seeks to minimize the maximum value of the PSF, W(a), within a specified interval (the clean field of view, or CLF), subject to constraints on the positions and moduli of the telescopes.

A narrow central lobe in the PSF indicates a higher angular resolution in the imaging system, which is desirable in optical astronomy. By optimizing the PSF with the goal of minimizing its maximum value within the CLF, you indirectly aim to achieve a narrow central lobe and low secondary lobes. This would result in better resolution and dynamic range for the imaging system.

It is important to note that minimizing the PSF in the optimization problem is not the same as minimizing the width of the central lobe directly. The optimization problem is formulated to find the best possible configuration of the telescope array, taking into account multiple factors, including the main lobe and secondary lobes of the PSF. The narrow central lobe is a desired outcome of this optimization process, but the optimization problem itself focuses on minimizing the maximum value of the PSF within the CLF.

Our optimization model is then

Minimize : Max{ $\Psi(\alpha)$: $\alpha_{min} \leq \alpha \leq \alpha_{max}$ }

Subject to : 1. $u_{k+1} - u_k \leq d$, for $k = 1,2,3,...(n-1)$,

2. $\Sigma_{k=1}^{n} a_k = 1$,

3. $a_k \geq 0$, for $k = 0,1,2,3,...,n$.

Now we can rephrase this problem more into a computational form by using above equations.

Minimize : t

Subject to : 1. $\left| \frac{J_1(\pi\alpha)}{\alpha} \Sigma_{k=1}^{n} a_k \, e^{-\frac{2i\pi u_k}{d}\alpha} \right| \leq t$, $\alpha_{min} \leq \alpha \leq \alpha_{max}$,

2. $u_{k+1} - u_k \leq d$, for $k = 1,2,3,...(n-1)$

3. $\Sigma_{k=1}^{n} a_k = 1$,

4. $a_k \geq 0$, for $k = 0,1,2,3,...,n$

This is a Semi-infinite nonlinear optimization problem. There is only a finite number of optimization variables, but an infinite number of constraints. To solve the problem we simply use a Discretization method.

Suppose now that there is an even number $(n = 2m)$ of pupils. The case of an odd number could be considered similarly. We assume also that the pupils are symmetrically placed around zero. By discretizing the interval $[\alpha_{min}, \alpha_{max}]$ with q points uniformly spaced, the optimization model becomes

Minimize : t

Subject to : 1. $-t \leq \frac{J_1(\pi\alpha_j)}{\alpha_j} \Sigma_{k=1}^{n} a_k \; cos(\frac{2\pi}{d} \; u_k \; \alpha_j) \leq t$ for $j = 1,2,...q, \; u_1 \geq \frac{d}{2}$,

2. $u_{k+1} - u_k \leq d$, for $k = 1,2,3,...(n-1)$

3. $\Sigma_{k=1}^{n} a_k = 1$,

4. $a_k \geq 0$, for $k = 0,1,2,3,...,n$

This is a nonlinear optimization problem with a finite number of variables and constraints.

What is a Semi – Infinite Non Linear Optimization Problem ?

A semi-infinite nonlinear optimization problem is a particular type of optimization problem where the objective function and/or constraints involve an infinite number of variables, or more specifically, the constraints are defined over an infinite set, usually a continuous domain. In this context, nonlinear refers to the fact that the objective function or constraints are not linear functions of the decision variables.

Solving semi-infinite nonlinear optimization problems can be challenging due to the infinite nature of the constraints. Typically, specialized algorithms and numerical techniques are employed to handle the infinite constraints and find an optimal solution. These methods may include discretization, cutting plane methods, or other techniques that approximate the infinite constraint set by a finite one.

What is Discretization Method and how it is going to use in our Optimization Problem ?

Discretization is a numerical method used to convert continuous functions, equations, or domains into a discrete form. It involves dividing the continuous domain into a finite number of smaller, discrete parts, such as intervals or grid points. Discretization is often used in solving complex optimization problems, partial differential equations, and other mathematical problems that involve continuous domains or functions.

In the context of an optimization problem, especially when dealing with semi-infinite or infinite constraints, discretization can be used to approximate the continuous constraint domain by a finite set of discrete points. By doing so, the problem can be transformed into a more manageable, finite-dimensional optimization problem that can be solved using standard optimization techniques.

I. Identify the continuous domain : Determine the continuous domain or constraint set that needs to be discretized.
II. Discretize the domain : Divide the continuous domain into a finite number of smaller, discrete parts. This can be done by creating a grid, partitioning the domain into intervals, or using other discretization techniques.
III. Reformulate the problem : Replace the continuous constraints with their discrete counterparts. This often involves replacing integrals with sums, continuous functions with discrete approximations, and so on.
IV. Solve the discretized problem : Use standard optimization techniques to solve the discretized problem.
V. Analyze the solution : Assess the quality of the solution obtained from the discretized problem. Depending on the discretization method and the specific problem, the solution may need to be refined or the discretization adjusted to obtain a more accurate approximation of the original continuous problem.

How do we solve the Non Linear Optimization Problem with finite number of variable which we found above ?

To solve the nonlinear optimization problem with a finite number of variables, you can use a variety of optimization algorithms and techniques. The choice of method depends on the specific characteristics of the problem, such as the presence of constraints, smoothness of the objective function, and the availability of gradient information.