

## Chapter 2

# Mathematics for Interval Algebra and Optimization

**Abstract** This chapter introduces some mathematical algorithms for solving interval linear equations, interval nonlinear equations, interval nonlinear equations and interval robust optimization models.

### 2.1 Definition of Intervals

**Definition:** An interval  $\langle x \rangle$  is a non-empty set of real numbers, satisfying  $\langle x \rangle = [\underline{x}, \bar{x}] = \{x \in \mathbb{R} : \underline{x} \leq x \leq \bar{x}\}$ , where  $\underline{x}$  and  $\bar{x}$  are called as the upper and lower bound of the interval. Moreover, we have  $\underline{x}, \bar{x} \in \mathbb{R}$  and  $\underline{x} \leq \bar{x}$ . In particular, when  $\underline{x} = \bar{x}$ , the interval is degenerated into a real number. The set of all intervals are called interval space, denoted by  $\mathbb{IR}$ .

Interval arithmetic operations are defined on interval space  $\mathbb{IR}$ . Given two intervals  $\langle x \rangle$  and  $\langle y \rangle$ , we have

Absolute value of intervals:  $|\langle x \rangle| \triangleq \text{mag}(\langle x \rangle) \triangleq \max\{|\underline{x}|, |\bar{x}|\}$ ;

Inclusion of intervals:  $\langle x \rangle \subseteq \langle y \rangle \triangleq \underline{y} \leq \underline{x} \text{ and } \bar{y} \geq \bar{x}$ ;

Basic arithmetic operations of intervals:

Addition  $\langle x \rangle + \langle y \rangle \triangleq [\underline{x} + \underline{y}, \bar{x} + \bar{y}]$ ;

Subtraction  $\langle x \rangle - \langle y \rangle \triangleq [\underline{x} - \bar{y}, \bar{x} - \underline{y}]$ ;

Multiplication  $\langle x \rangle \times \langle y \rangle \triangleq [\min\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\bar{y}, \bar{x}\underline{y}, \bar{x}\bar{y}\}]$ ;

Division  $\langle x \rangle \div \langle y \rangle \triangleq [1/\bar{y}, 1/\underline{y}] \times \langle x \rangle$ , where  $0 \notin \langle y \rangle$

Intersection of intervals:  $\langle x \rangle \cap \langle y \rangle = [\max(\underline{x}, \underline{y}), \min(\bar{x}, \bar{y})]$ ;

Union of intervals:  $\langle x \rangle \cup \langle y \rangle = [\min(\underline{x}, \underline{y}), \max(\bar{x}, \bar{y})]$ ;

Comparison of intervals:  $\langle x \rangle < \langle y \rangle \Leftrightarrow \bar{x} < \underline{y}$ ;

The interval length:  $\text{length}(\langle x \rangle) \triangleq \bar{x} - \underline{x}$ ;

For an interval vector  $\langle \mathbf{x} \rangle \in \mathbb{IR}^{n \times 1}$ , a norm of an interval vector is given by  $\|\langle \mathbf{x} \rangle\| = \max\{|\langle x_i \rangle| : i = 1, \dots, n\}$ . According to the definition of norm followed by (i)–(iv), we will show that  $\|\langle \mathbf{x} \rangle\|$  defined above is a norm of  $\langle \mathbf{x} \rangle$ :

- (i) Subadditivity:  $\|\langle \mathbf{x} \rangle\| = \max\{|\langle \mathbf{x}_i \rangle| : i = 1, \dots, n\} \geq 0 \Rightarrow \|\langle \mathbf{x} \rangle\| \geq 0$ ;
- (ii) Existence of zeros:  $\|\langle \mathbf{x} \rangle\| = 0 \Rightarrow \max(|\bar{\mathbf{x}}|, |\underline{\mathbf{x}}|) = 0 \Rightarrow |\bar{\mathbf{x}}| = |\underline{\mathbf{x}}| = 0 \Rightarrow \langle \mathbf{x} \rangle = [0, 0]$ ;
- (iii) Homogeneity:  $\|c \cdot \langle \mathbf{x} \rangle\| = \max(|c\bar{\mathbf{x}}|, |c\underline{\mathbf{x}}|) = |c|\max(|\bar{\mathbf{x}}|, |\underline{\mathbf{x}}|) = |c|\|\langle \mathbf{x} \rangle\|$ ;  
 $\|\langle \mathbf{x} \rangle + \langle \mathbf{y} \rangle\| = \max(|\bar{\mathbf{x}} + \bar{\mathbf{y}}|, |\underline{\mathbf{x}} + \underline{\mathbf{y}}|) \leq \max(|\bar{\mathbf{x}}| + |\bar{\mathbf{y}}|, |\underline{\mathbf{x}}| + |\underline{\mathbf{y}}|)$
- (iv) Positivity:  
 $\leq \max(|\bar{\mathbf{x}}|, |\underline{\mathbf{x}}|) + \max(|\bar{\mathbf{y}}|, |\underline{\mathbf{y}}|) = \|\langle \mathbf{x} \rangle\| + \|\langle \mathbf{y} \rangle\|$

The aforementioned definitions of intervals are simple and understood, however, it needs to know that the interval space is not satisfied with the definition of linear space.

Example 1:  $\langle \mathbf{x} \rangle - \langle \mathbf{x} \rangle = [2\underline{\mathbf{x}}, 2\bar{\mathbf{x}}] \neq 0$ . It shows that interval space is not a closure for subtraction.

Example 2:  $\langle \mathbf{x} \rangle(\langle \mathbf{y} \rangle + \langle \mathbf{z} \rangle) \neq \langle \mathbf{x} \rangle\langle \mathbf{y} \rangle + \langle \mathbf{x} \rangle\langle \mathbf{z} \rangle$ , instead  $\langle \mathbf{x} \rangle(\langle \mathbf{y} \rangle + \langle \mathbf{z} \rangle) \subseteq \langle \mathbf{x} \rangle\langle \mathbf{y} \rangle + \langle \mathbf{x} \rangle\langle \mathbf{z} \rangle$ . It shows that the distributive law does not hold in interval space.

The two examples above show interval space does not follow operational rules of linear space, such that the interval solution will be calculated greater and greater, especially when the same intervals appear several times in one expression, the range of interval will become conservative. Therefore, it is significant to study how to reduce the conservatism of interval arithmetic.

Furthermore, define two special interval matrices (M-matrix and H-matrix) and a special real matrix (comparison matrix). Given an interval matrix  $\langle \mathbf{A} \rangle \in \mathbb{IR}^{n \times n}$ ,  $\langle \mathbf{A} \rangle$  is called an M-matrix if and only if  $\langle \mathbf{A} \rangle_{ij} \leq 0$  for all  $i \neq j$  and  $\langle \mathbf{A} \rangle \mathbf{u} > 0$  for some positive vector  $\mathbf{u} \in \mathbb{R}^n$ ; the comparison matrix  $\mathbf{A}^c$  of  $\langle \mathbf{A} \rangle$  is defined as (2.1), if  $\mathbf{A}^c$  is an M-matrix, then  $\langle \mathbf{A} \rangle$  is said to be an H-matrix. Additionally, the central matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  of interval matrix  $\langle \mathbf{A} \rangle \in \mathbb{IR}^{n \times n}$  is defined as: for  $\forall [a_{ij}] \in \langle \mathbf{A} \rangle$ ,  $a_{ij} = (\bar{a}_{ij} + \underline{a}_{ij})/2 \in \mathbf{A}$

$$\mathbf{A}^c = \begin{cases} \text{mig}(\langle a_{ii} \rangle) & i = j \\ -\text{mag}(\langle a_{ij} \rangle) & i \neq j \end{cases} \quad \forall [a_{ij}] \in \mathbf{A} \quad (2.1)$$

## 2.2 Solutions of Algebraic Equations with Right-Hand Intervals

### 2.2.1 Linear Equations with Right-Hand Intervals Using Kraw Operator Iteration Method with Preconditioner

Given  $\langle \mathbf{A} \rangle \in \mathbb{IR}^{n \times n}$  and  $\langle \mathbf{b} \rangle \in \mathbb{IR}^{n \times 1}$ , traditional interval linear equations are defined as  $\langle \mathbf{A} \rangle \langle \mathbf{x} \rangle = \langle \mathbf{b} \rangle$  to solve  $\langle \mathbf{x} \rangle$ . To solve general interval linear equations, four solution methods are discussed in the paper including interval Gauss elimination

method, Kraw operator iteration method, interval hull method and optimality-based method. The details and flow chart of aforementioned four methods can be found in Appendix A.

However, if the coefficient matrix of interval linear equations is a constant matrix, the traditional interval linear equations can be simplified as linear equations with right-hand intervals, such that  $\mathbf{A}\langle\mathbf{x}\rangle = \langle\mathbf{b}\rangle$ . Thus, the special structural interval linear equations obviously can be solved by four traditional methods:

- (i) For interval Gauss elimination method, interval operations of coefficient matrix  $\langle\mathbf{A}\rangle$  in the method directly replace by operations of traditional deterministic coefficient matrix  $\mathbf{A}$  in the real space, and right-hand intervals still adopt interval operations. However, the nature of interval Gauss elimination method is a direct method, and time complexity of the method is excessive.
- (ii) For Kraw operator iteration method and interval hull method, they both need to compute and express the inverse of central matrix  $(\mathbf{A}^c)^{-1}$ . However, for large system, it is difficult to directly compute the inverse of the matrix.
- (iii) For optimality-based method, it is effective to solve this problem, because interval linear equations can be transformed into linear programming followed by (2.2)–(2.3). The linear programming is easy to solve, but we can find that the interval of every variable need to compute separately. Certainly, the interval of every variable is independent and can use parallel implementation. However, if the number of variables are relatively large, it will need large parallel cores to implement while (i) and (ii) can solve intervals of all variables simultaneously.

$$\min/\max_{x, \mathbf{A}} x_i, \quad i = 1, \dots, n \quad (2.2)$$

$$\underline{b}_j \leq \sum_{i=1}^n a_{ij}x_i \leq \overline{b}_j, \quad j = 1, \dots, n \quad (2.3)$$

With respect to conservatism problem, because interval operations are not satisfied with distributive law, LU decomposition will also expand the range of the interval solution even for interval linear equations with constant coefficient. To reduce the conservatism of the interval arithmetic, four methods in Appendix A can overcome conservatism for interval linear equations with constant coefficient. Therefore, the high efficiency of solution method should be focused on.

Reviewing the four methods, interval hull method needs to solve inverse central matrix of coefficient matrix, so the complexity of computation is more difficult than other methods; the relatively direct method (interval Gauss elimination method) and the general iteration method (Kraw operator iteration method) has a faster solution speed and a lower space complexity, but iteration method usually has poor

convergence. Kraw operator iteration method needs to solve inverse of central matrix to guarantee the convergence. However, the book to compute inverse of a matrix is much great, and it hard to implement for a large system in practice. According to the iteration expression of Kraw operator iteration method, we only need to find a better matrix which is defined to guarantee the convergence, so krawczyk iteration method with approximate inverse preconditioner can be used [1, 2]. For coefficient matrix  $A$ , a state-of-the-art method is to minimize the Frobenius norm of the residual matrix  $\|I - CA\|_F^2$  to seek an approximate inverse  $C$ . Generally, we can take the 2-norm into consideration. In addition, the 2-norm can be decoupled as the sum of the squares, which can be utilized for the parallel computation. The optimization model is formulated as following:

$$\min_C \|I - CA\|_2^2 = \sum_{i=1}^n \|e_i - C_i A\|_2^2 \quad (2.4)$$

where  $C_i^{k+1}$  denotes  $i$ th row of  $C$  at  $k + 1$  iteration,  $n$  is the dimension of coefficient matrix  $A$ , and  $e_i$  is the  $i$ th row of the identity matrix.

Furthermore, an approximate inverse can be computed using Newton's method with the iteration as

$$C^{k+1} = C^k - (I - C^k A)(-C^k) = C^k(2I - AC^k) \quad (2.5)$$

$$C_i^{k+1} = C_i^k - (e_i - C_i^k A)(-C^k) = C_i^k(2I - AC^k), \quad i = 1, \dots, n \quad (2.6)$$

**Theorem 2.1** *To guarantee the convergence of Kraw operator iteration method, constant matrix  $C$  should satisfy  $\|I - CA\| = \beta < 1$*

*Proof* According to Kraw operator iteration expression, define an interval sequence  $\{\langle x^k \rangle\}$  which satisfies  $\langle x^{k+1} \rangle = (C\langle b \rangle + (I - CA)\langle x^k \rangle) \cap \langle x^k \rangle \subseteq (C\langle b \rangle + (I - CA)\langle x^k \rangle)$ .

Then, we introduce a new sequence  $\{\langle y^k \rangle\}$  with  $\langle y^{k+1} \rangle = (C\langle b \rangle + (I - CA)\langle y^k \rangle)$ . We can easily derive that  $0 \leq \|\langle x^k \rangle\| \leq \|\langle y^k \rangle\|$ , so if  $\{\langle y^k \rangle\}$  converges,  $\{\langle x^k \rangle\}$  will be bounded by the limit of  $\{\langle y^\infty \rangle\}$ .

For the sequence  $\{\langle y^k \rangle\}$ , it derives

$$\begin{cases} \langle y^{k+1} \rangle = (C\langle b \rangle + (I - CA)\langle y^k \rangle), & \|\langle y^{k+1} \rangle\| = \|C\langle b \rangle + (I - CA)\langle y^k \rangle\| \\ \langle y^k \rangle = (C\langle b \rangle + (I - CA)\langle y^{k-1} \rangle), & \|\langle y^k \rangle\| = \|C\langle b \rangle + (I - CA)\langle y^{k-1} \rangle\| \end{cases} \quad (2.7)$$

Take the difference of the two equations, and we have

$$\|\langle y^{k+1} \rangle - \langle y^k \rangle\| \leq \lambda \|\langle y^k \rangle - \langle y^{k-1} \rangle\| \leq \lambda^2 \|\langle y^{k-1} \rangle - \langle y^{k-2} \rangle\| \leq \dots \leq \lambda^k \|\langle y^1 \rangle - \langle y^0 \rangle\| \quad (2.8)$$

Next, we will show that if  $n, m \geq N$  for an any given number  $N$  (assuming  $n > m$ ), then it holds that

$$\|\langle \mathbf{y}^n \rangle - \langle \mathbf{y}^m \rangle\| \leq \frac{\lambda^n}{1 - \lambda} \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \quad (2.9)$$

It can be proved as following:

$$\begin{aligned} \|\langle \mathbf{y}^{m+1} \rangle - \langle \mathbf{y}^m \rangle\| &\leq \lambda^m \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \\ \|\langle \mathbf{y}^m \rangle - \langle \mathbf{y}^{m+2} \rangle\| &\leq \|\langle \mathbf{y}^{m+2} \rangle - \langle \mathbf{y}^m \rangle\| \leq \|\langle \mathbf{y}^{m+1} \rangle - \langle \mathbf{y}^m \rangle\| \\ &\quad + \|\langle \mathbf{y}^{m+2} \rangle - \langle \mathbf{y}^{m+1} \rangle\| \leq (\lambda^m + \lambda^{m+1}) \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \dots \\ \|\langle \mathbf{y}^n \rangle - \langle \mathbf{y}^m \rangle\| &\leq (\lambda^m + \lambda^{m+1} + \dots + \lambda^{n-1}) \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \\ &= \lambda^m (1 + \lambda + \dots + \lambda^{n-m-1}) \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \\ &\leq \lambda^n (1 + \lambda + \lambda^2 + \dots) \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| = \frac{\lambda^n}{1 - \lambda} \|\langle \mathbf{y}^1 \rangle - \langle \mathbf{y}^0 \rangle\| \end{aligned}$$

(Q.E.D.)

Thus, we can conclude that if  $\lambda < 1$ ,  $\{\langle \mathbf{y}^k \rangle\}$  is a Cauchy sequence, which converges to a point  $\langle \mathbf{y}^* \rangle$  that belongs to  $\mathbb{IR}$  space. Thus,  $\{\langle \mathbf{x}^k \rangle\}$  will be bound by  $\{\langle \mathbf{y}^* \rangle\}$ , such that  $0 \leq \|\langle \mathbf{x} \rangle\| \leq \|\langle \mathbf{y}^* \rangle\|$ .

On the other hand

$$\langle \mathbf{x}^{k+1} \rangle = (\mathbf{C}\langle \mathbf{b} \rangle + (\mathbf{I} - \mathbf{CA})\langle \mathbf{x}^k \rangle) \cap \langle \mathbf{x}^k \rangle \subseteq \langle \mathbf{x}^k \rangle \Rightarrow \|\langle \mathbf{x}^{k+1} \rangle\| \leq \|\langle \mathbf{x}^k \rangle\|$$

As a result, we can see that  $\{\langle \mathbf{x}^k \rangle\}$  is a monotonously decreasing and bounded sequence; therefore, it must be converged after several iterations.

(Q.E.D.)

According to Corollary 2.1, the spectral radius of  $\mathbf{I} - \mathbf{C}_0\mathbf{A}$  must be less than one with the aim of convergence. Therefore, if the initial guess of  $\mathbf{C}^0$  is chosen as  $\mathbf{C}^0 = \delta\mathbf{A}^T$ , it yields

$$\rho(\mathbf{I} - \delta\mathbf{A}^T\mathbf{A}) < 1 \Rightarrow 0 \leq \delta \leq \frac{2}{\rho(\mathbf{A}^T\mathbf{A})} \quad (2.10)$$

where  $\rho$  denotes the spectral radius of a matrix.

It is easily known from [3] that  $\rho(\mathbf{A}^T\mathbf{A}) \leq \|\mathbf{A}^T\mathbf{A}\|$ , where  $\|\cdot\|$  is any subordinate norm. In practice, we use

$$\delta = 1/\|\mathbf{A}^T\mathbf{A}\|_1 \quad (2.11)$$

Finally, it is important to point out that the approximate inverse matrix may become denser and denser as the iterations progress, which will need additional

communication with memory and will have a side effect on the computation efficiency. Meanwhile, some nonzero elements with small value is extremely close to zero, but they can not be fully equal to zero because of calculation error. Therefore, the sparsity representation should be performed during each iteration by dropping some elements with small value in the approximate inverse matrix. The flowchart of proposed Kraw operator iteration method with minimum norm preconditioner is shown in Table 2.1.

**Table 2.1** The flowchart of Kraw operator iteration method with minimum norm preconditioner

Approximate Inverse Preconditioner-based Kraw Operator Iteration Method	
<b>Input:</b>	$\varepsilon$ is a preset convergence precision, $n$ is the dimension of coefficient matrix $A$ , and $\eta$ is the dropping tolerance.
<b>Output:</b>	The interval solution $\langle x \rangle$ of $A \langle x \rangle = \langle b \rangle$
<b>Step 1:</b>	Compute $\delta$ by (2-11), such that $\delta = 1/\ A^T A\ $ ;
<b>Step 2:</b>	Let $C^0 \leftarrow \delta A^T$ , $\lambda \leftarrow +\infty$ , and $k \leftarrow 1$ ;
<b>Step 3:</b>	<b>while</b> $\lambda > 1$
<b>Step 4:</b>	<b>for</b> $i=1:n$
<b>Step 5:</b>	$TempC_i^k \leftarrow C_i^{k-1}(2I - AC^k)$ ;
<b>Step 6:</b>	$TCmax_i^k \leftarrow \max(TempC_i^k)$ ;
<b>Step 7:</b>	<b>for</b> $j=1:n$
<b>Step 8:</b>	<b>if</b> $TempC_{i,j}^k/TCmax_i^k \leq \eta$
<b>Step 9:</b>	$TempC_{i,j}^k \leftarrow 0$ ;
<b>Step 10:</b>	<b>end</b>
<b>Step 11:</b>	<b>end</b>
<b>Step 12:</b>	<b>end</b>
<b>Step 13:</b>	$C^k \leftarrow TempC^k$ and $k \leftarrow k+1$ ;
<b>Step 14:</b>	$\lambda \leftarrow \ I - C^k A\ _2$ ;
<b>Step 15:</b>	<b>end</b>
<b>Step 16:</b>	Let $b \leftarrow Cb$ , $A \leftarrow I - CA$ ;
<b>Step 17:</b>	Compute $\lambda \leftarrow \ A\ $ and $\alpha = \ C \langle b \rangle\ /(1-\lambda)$ ;
<b>Step 18:</b>	Let $\langle x^0 \rangle \leftarrow ([-\alpha, +\alpha], [-\alpha, +\alpha], \dots, [-\alpha, +\alpha])^T$ , $s^0 \leftarrow 2n\alpha$ , $\tau \leftarrow +\infty$ and $k \leftarrow 1$ ;
<b>Step 19:</b>	<b>while</b> $\tau \leq \varepsilon$
<b>Step 20:</b>	$\langle x^{k+1} \rangle \leftarrow (C \langle b \rangle + (I - CA) \langle x^k \rangle) \cap \langle x^k \rangle$
<b>Step 21:</b>	$s^{k+1} \leftarrow \text{sum}(\text{length}(\langle x^{k+1} \rangle))$ ;
<b>Step 22:</b>	$\tau \leftarrow (s^k - s^{k+1})/s^{k+1}$ ;
<b>Step 23:</b>	<b>end</b>

### 2.2.2 Nonlinear Equations with Right-Hand Intervals Using Optimality-Based Method

It should be noted that the interval nonlinear equations are more difficult than interval linear equations to be handled. If the intervals only exist in the right hand of the nonlinear equations, the interval nonlinear equations can be compactly formulated as  $F(\langle x \rangle) = \langle p \rangle$ , where  $p$  is the input interval parameters and  $\langle x \rangle$  is the interval solution of the nonlinear equations. Generally, the interval nonlinear equations can be solved by the interval nonlinear Kraw operator iteration method, the flowchart of which can be found in the Appendix B.

However, the conservatism problem of interval nonlinear equations are more serious. As Sect. 1.2.1 mentioned, when the same intervals appear several times in one expression, the range of interval will become conservative. However, in the computation of interval nonlinear equations, it may have many same elements in Jacobian matrix and initial equations, so the interval solutions may be greatly expanded after several iterations, which will lead to serious conservatism problem. Therefore, the biggest challenge to solve interval nonlinear equations is to overcome conservatism of interval operations.

Similarly, optimality-based method of interval linear equations can be used. If  $p$  is regarded as a new variable, then in a high-dimensional space with the optimal variable of  $(p, x)$ , the optimization model is formulated as following:

$$\min/\max_{x,p} \quad x_i \quad s.t. \quad \underline{p} \leq F(x, p) \leq \bar{p}, \quad i = 1, \dots, n \quad (2.12)$$

With the objective function of min and max respectively, the lower and upper bound of the variable  $i$  can be solved. Similarly, because the solution of every variable is independent, it can be divided into  $n$  optimizer with parallel implementation. Compared with Kraw operator iteration method, optimality-based method is able to attain compact range of interval. However, the aforementioned model is generally a nonlinear optimization model, and it is hard to compute global optimal solution. Then, our discussion is based on the common interval quadratic nonlinear equations.

#### 2.2.2.1 Linear Relaxation Model for Interval Quadratic Equations with Optimality-Based Method

Consider interval quadratic nonlinear equations  $x^T Q_j x + c_j^T x = [d_j, \bar{d}_j], j = 1, \dots, m$ . The optimization model based the aforementioned method is formulated as

$$\max/\min_{x \in \mathbb{R}^{n \times 1}} \quad x^T Q_0 x + c_0^T x \quad (2.13)$$

$$s.t. \quad \underline{d}_j \leq \mathbf{x}^T \mathbf{Q}_j \mathbf{x} + \mathbf{c}_j^T \mathbf{x} \leq \overline{d}_j, \quad j = 1, \dots, m \quad (2.14)$$

It notes that the independent variable  $x$  generally has a reasonable given range of interval. Let the range of interval is  $l_i \leq x_i \leq u_i, i = 1, \dots, n$ , then we can get a general quadratically constrained quadratic programming (QCQP) model:

$$(\text{QCQP}) \quad \max/\min_{\mathbf{x} \in \mathbb{R}^{n \times 1}} \quad \mathbf{x}^T \mathbf{Q}_0 \mathbf{x} + \mathbf{c}_0^T \mathbf{x} \quad (2.15)$$

$$s.t. \quad \underline{d}_j \leq \mathbf{x}^T \mathbf{Q}_j \mathbf{x} + \mathbf{c}_j^T \mathbf{x} \leq \overline{d}_j, \quad j = 1, \dots, m \quad (2.16)$$

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n \quad (2.17)$$

where  $n$  denotes the number of quadratic constraints,  $\mathbf{Q}_j$  ( $j = 1, \dots, m$ ) are indefinite  $n \times n$  matrices,  $\mathbf{c}_j$  ( $j = 0, \dots, m$ ) are  $n$ -dimensional vectors, the set  $\bigcup_{i=1}^n [l_i, u_i]$  is assumed to be nonempty and bounded.

$\phi(x, y)$  is called bilinear function if  $\phi(x, \cdot)$  and  $\phi(\cdot, y)$  are both linear functions. For example,  $\phi(x, y) = xy$  is a bilinear function. Let a mapping  $f: \Omega \rightarrow \mathbb{R}$  be a bilinear function, then  $\text{vex}_\Omega(f)$  is the convex envelope of  $f$  over  $\Omega$ , if it is the point-wise supremum of convex underestimation of  $f$  over  $\Omega$ . Similarly,  $\text{cav}_\Omega(f)$  is the concave envelope of  $f$  over  $\Omega$ , if it is the point-wise infimum of concave overestimation of  $f$  over  $\Omega$ , then a theorem is given as following.

**Theorem 2.2:**  $xy$  is a bilinear function given in a bounded closed set  $\Psi$ . The convex envelope  $\text{vex}_\Omega(xy)$  and concave envelope  $\text{cav}_\Omega(xy)$  are determined by (2.58) and (2.59), and then we can arrive at  $\text{vex}_\Omega(xy) \leq xy \leq \text{cav}_\Omega(xy)$ .

$$\Psi \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2 \mid l_x \leq x \leq u_x, l_y \leq y \leq u_y\} \quad (2.18)$$

$$\text{vex}_\Omega(xy) = \max\{l_y x + l_x y - l_x l_y, u_y x + u_x y - u_x u_y\} \quad (2.19)$$

$$\text{cav}_\Omega(xy) = \min\{u_y x + l_x y - l_x u_y, l_y x + u_x y - u_x l_y\} \quad (2.20)$$

Let dummy variable  $\mathbf{Z} = \mathbf{x}\mathbf{x}^T \in \mathbb{R}^{n \times n}$  and substitute  $\mathbf{Z}$  for  $\mathbf{x}\mathbf{x}^T$  in QCQP model, the objective and constraints become bilinear. Furthermore, the convex envelopes and concave envelopes as defined in Theorem 2.2 are introduced to obtain the tractable linear relaxation (LR) model of the non-convex problem QCQP, formulated as follows:

$$(\text{QCQP-LR}) \quad \max/\min_{\mathbf{x} \in \mathbb{R}^{n \times n}, \mathbf{z} \in \mathbb{R}^{n \times n}} \quad \mathbf{c}^T \mathbf{x} + \mathbf{Q}_0 \cdot \mathbf{Z} \quad (2.21)$$

$$s.t. \quad \underline{d}_j \leq \mathbf{c}_j^T \mathbf{x} + \mathbf{Q}_j \cdot \mathbf{Z} \leq \overline{d}_j, \quad j = 1, \dots, m \quad (2.22)$$



$$z_{ij} - l_i x_j - l_j x_i + l_i l_j \geq 0, \quad j = 1, \dots, m \quad i = 1, \dots, n \quad (2.23)$$

$$z_{ij} - u_i x_j - u_j x_i + u_i u_j \geq 0, \quad j = 1, \dots, m \quad i = 1, \dots, n \quad (2.24)$$

$$z_{ij} - l_i x_j - u_j x_i + l_i u_j \leq 0, \quad j = 1, \dots, m \quad i = 1, \dots, n \quad (2.25)$$

$$z_{ij} - u_i x_j - l_j x_i + u_i l_j \leq 0, \quad j = 1, \dots, m \quad i = 1, \dots, n \quad (2.26)$$

$$l_i \leq x_i \leq u_i, \quad i = 1, \dots, n \quad (2.27)$$

where  $\forall A, B \in \mathbb{R}^{n \times n}$ ,  $A \cdot B = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij}$ .

For minimization problem of the original QCQP, the QCQP-LR is a relaxed linear programming model and therefore provides a lower bound of the original QCQP. Similarly, it provides an upper bound for the maximization problem of the original QCQP.

### 2.2.3 Optimality-Based Bounds Tightening Method

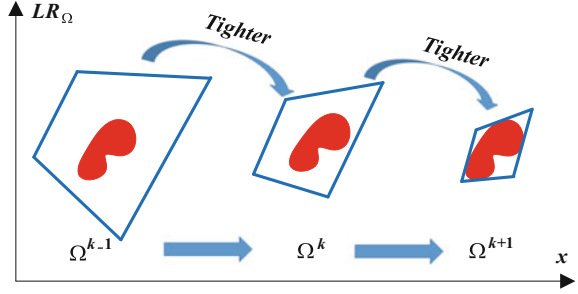
It can be observed that the relaxed QCQP-LR model is dependent on the hyper-rectangular variable domain. At the beginning, the initial domain is usually so large that leads to a conservational interval solution, so it is desirable to schedule a more rigorous hyper-rectangle domain to achieve an optimal interval solution. For this reason, the optimality-based bounds tightening (OBBT) method proposed in [4–6] is employed to solve the QCQP-LR model, which will tighten the relaxed model QCQP-LR based on linear optimality by cycling through each participating variable until the volume fails to improve.

Define  $\mathbf{x}^k = (\mathbf{e}^k, \mathbf{f}^k, \mathbf{y}^k)$  to represent the variable set in the  $k$ th iteration. Let  $LR_{\Omega}^k(\mathbf{x}^k)$  denote the constraints of QCQP-LR model and let  $(\varphi_{\Omega})^k(\cdot)$  be the optimal objective value of (QCQP-LR) in the feasible domain  $\Omega^k = \{\underline{\mathbf{x}}^k \leq \mathbf{x} \leq \overline{\mathbf{x}}^k\}$ . Then, the hyper-rectangular variable domain can be tightened or ‘squeezed’ through an iterative linear programming solver, which is illustrated in the following three steps.

- (a) The initial matrix domain is set as  $\Omega^0$ , set  $k \leftarrow 0$ ;
- (b) Solve  $\begin{cases} \bar{x}_i^{k+1} \leftarrow \{\max x_i \text{ s.t. } LR_{\Omega^k}(\mathbf{x}^k) \cap \Omega^k\} \\ \underline{x}_i^{k+1} \leftarrow \{\min x_i \text{ s.t. } LR_{\Omega^k}(\mathbf{x}^k) \cap \Omega^k\} \end{cases}$  for  $\forall I$ , set  $k \leftarrow k + 1$ ;
- (c)  $\Omega^k \leftarrow \{\underline{\mathbf{x}}^k \leq \mathbf{x} \leq \overline{\mathbf{x}}^k\}$ , if  $\Omega^k = \Omega^{k-1}$  stop; otherwise, go to (b).

From iteration process above, we know that interval power flow computation includes ‘min’ and ‘max’ models simultaneously solved. Based on OBBT method, the variable domain  $\Omega^i$  satisfies  $\Omega^{i-1} \subseteq \Omega^i$ , and as such the domain sequence  $\{\Omega^0, \Omega^1, \Omega^2, \dots, \Omega^i, \dots\}$ , obtained from step (b), is monotonously decreasing and geometrically a nested sequence of hyper-rectangles, so the domain sequence is a

**Fig. 2.1** The process of gradually tightened variable domain using OBBT-based method



Cauchy sequence and has a limit. Additionally, QCQP-LR model is a linear programming in a bounded closed set (compact set). For any given solution  $\mathbf{x}$  and  $\mathbf{y}$ , we have  $\|(\varphi_\Omega)(\mathbf{x}) - (\varphi_\Omega)(\mathbf{y})\|_\infty = \|(\mathbf{x}^T \mathbf{Q}_0 \mathbf{x} + c_0^T \mathbf{x}) - (\mathbf{y}^T \mathbf{Q}_0 \mathbf{y} + c_0^T \mathbf{y})\|_\infty \leq \|\mathbf{u} - \mathbf{l}\|_\infty \cdot \|\mathbf{Q}_0\|_\infty \cdot \|\mathbf{u} - \mathbf{l}\|_\infty < +\infty$ , which means operator  $(\varphi_\Omega)^i(\cdot)$  is a bounded operator. Therefore,  $\|(\varphi_\Omega)^{i+1}(\mathbf{x}) - (\varphi_\Omega)^i(\mathbf{x})\|_\infty \leq \|(\varphi_\Omega)^i(\mathbf{x})\|_\infty \cdot \|\Omega^{i+1} - \Omega^i\|_\infty$ , where  $\|\Omega^i\|_\infty = \|\bar{\mathbf{x}}^i - \underline{\mathbf{x}}^i\|_\infty$ . According to functional analysis, the aforementioned expression denotes that a Cauchy sequence is mapped to a Cauchy sequence by bounded operator. Therefore, the sequence  $\{(\varphi_\Omega)^0(\mathbf{x}), \dots, (\varphi_\Omega)^i(\mathbf{x}), \dots\}$  is also a Cauchy sequence and has a limit, which demonstrates that the proposed OBBT-based algorithm can converge after finite number of iterations. Furthermore, the process of gradually tightened hyper-rectangular variable domain is illustrated in Fig. 2.1, where the solid red area represents the feasible region of the original non-convex QCQP problem.

After the process stops (and OBBT algorithm converges), the optimally tightened hyper-rectangular variable domain  $\Omega^*$  and the corresponding interval solution are obtained.

## 2.3 Optimization Solutions Based Intervals

### 2.3.1 Optimization Solutions with Right-Hand Intervals

Interval linear optimization model with right-hand interval uncertainties is formulated as

$$\text{Min } Z = \mathbf{c}^T \mathbf{x} \quad (2.28)$$

$$\text{s.t. } \mathbf{A}\mathbf{x} = [\mathbf{w}^-, \mathbf{w}^+] \quad (2.29)$$

$$\mathbf{B}\mathbf{x} \leq [\mathbf{d}^-, \mathbf{d}^+] \quad (2.29)$$

$$l \leq x \leq u \quad (2.30)$$

where,  $A \in \mathbb{R}^{s \times n}$ ,  $B \in \mathbb{R}^{m \times n}$ ,  $c \in \mathbb{R}^{n \times 1}$ ,  $x \in \mathbb{R}^{n \times 1}$ ;  $m$  represents the number of interval inequations,  $s$  represents the number of interval equations,  $n$  represents the dimension of variable.

The objective of interval optimization is to find out the optimistic value and pessimistic value of objective functions and range of interval variables. For  $\forall w \in [w^-, w^+]$ ,  $\forall d \in [d^-, d^+]$ , solve the deterministic optimization model as following, and obtain the objective value and the optimal solution. if uncertainty variables are taken as any possible combinations of intervals, then we can obtain any possible objective values and optimal interval solutions, and the upper and lower bounds of which are called the optimal interval function value and the optimal interval solution, where the upper bound of optimal interval function value is called pessimistic value and the lower bound is called optimistic value.

$$\text{Min } Z = c^T x \quad (2.31)$$

$$s.t. \quad Ax = w \quad (2.32)$$

$$Bx \leq d \quad (2.33)$$

$$l \leq x \leq u \quad (2.34)$$

We can see that independent variable  $x$  and objective value  $Z$  of model (2.28)–(2.30) is the function of uncertainties, which is an optimization model. From this perspective, interval optimization is a special kind of interval mathematical operation, which is hard to obtain an explicit expression. Considering to overcome the conservatism of interval operations, it can be solved by the optimization method. Therefore, interval optimization can be regarded as the optimization of optimization, that is, bi-level programming.

Firstly, assume the interval optimization model is solvable, such that for  $\forall w \in [w^-, w^+]$ ,  $\forall d \in [d^-, d^+]$ , the optimization model (2.31)–(2.34) is feasible. Based on the assumption, the solution models of optimistic value and pessimistic value are obtained according to the two standard bi-level optimization model (2.35)–(2.39) and (2.40)–(2.44), where optimistic value denotes the optimal value among the range of uncertain disturbance interval, which is expressed as “min-min” form while pessimistic value denotes the worst value among the range of uncertain disturbance interval, which is expressed as “max-min” form

$$(\text{Optimistic value}) \min_{w,d} \min_x Z^- = c^T x \quad (2.35)$$

$$s.t. \quad Ax = w \quad (2.36)$$

$$\mathbf{B}\mathbf{x} \leq \mathbf{d} \quad (2.37)$$

$$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \quad (2.38)$$

$$\forall \mathbf{w} \in [\mathbf{w}^-, \mathbf{w}^+], \quad \forall \mathbf{d} \in [\mathbf{d}^-, \mathbf{d}^+] \quad (2.39)$$

$$\text{(Pessimistic value)} \quad \max_{\mathbf{w}, \mathbf{d}} \min_{\mathbf{x}} \quad Z^+ = \mathbf{c}^T \mathbf{x} \quad (2.40)$$

$$s.t. \quad \mathbf{A}\mathbf{x} = \mathbf{w} \quad (2.41)$$

$$\mathbf{B}\mathbf{x} \leq \mathbf{d} \quad (2.42)$$

$$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \quad (2.43)$$

$$\forall \mathbf{w} \in [\mathbf{w}^-, \mathbf{w}^+], \quad \forall \mathbf{d} \in [\mathbf{d}^-, \mathbf{d}^+] \quad (2.44)$$

#### A. The solution of optimistic model

Because the inner and outer level of bi-level optimistic model are both “min” problems, it can be combined as single-level programming, and we hope the feasible region is as large as possible to guarantee the lower bound of interval solution. Then, the inequality constraint is relaxed to the maximum bound  $\mathbf{d}^+$ , and the simplified single-level optimization is formulated as (2.45)–(2.48). Therefore, the linear programming problem can be solved effectively by simplex method.

$$\text{(Optimistic value)} \quad \min_{\mathbf{x}} \quad Z^- = \mathbf{c}^T \mathbf{x} \quad (2.45)$$

$$s.t. \quad \mathbf{w}^- \leq \mathbf{A}\mathbf{x} \leq \mathbf{w}^+ \quad (2.46)$$

$$\mathbf{B}\mathbf{x} \leq \mathbf{d}^+ \quad (2.47)$$

$$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \quad (2.48)$$

#### B. The solution of pessimistic model

However, for bi-level pessimistic model, the inner level is “min” model while the outer level is “max” model. Because the objectives of inner and outer levels are different, it can not be transformed as a single-level programming. Generally, it is easy to handle interval inequality constraints and hard to handle interval equality constraints. From the view of mathematics, strong duality theorem is employed to express the inner model as its dual model and a “max” model is obtained. Then the “max-max” model is transformed into a single-level “max” model according to the solution of two-level optimistic model. To pessimistic model, we hope the feasible region is as small as possible, so the inequality constraint is shrunk to the

minimum bound  $d^-$ . In the inner model, the uncertain value is regarded as a constant and dual model is used, then we have,

$$\begin{aligned}
 Z^+ &= \text{Max}_{b \leq w \leq b} \left\{ \begin{array}{l} \text{Min}_x \quad Z = c^T x \\ \text{s.t.} \quad Bx \leq d^- \quad \text{the number of constraints is } m \\ Ax = w \quad \text{the number of constraints is } s \\ Ix \geq l \quad \text{the number of constraints is } n \\ Ix \geq u \quad \text{the number of constraints is } n \end{array} \right. \\
 &\Leftrightarrow \\
 Z^+ &= \text{Max}_{b \leq w \leq b} \left\{ \begin{array}{l} \text{Max}_{y_1, y_2, y_3, y_4} \quad (d^-)^T y_1 + w^T y_2 + l^T y_3 + u^T y_4 \\ \text{s.t.} \quad \begin{bmatrix} E^T, A^T, I, I \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = c \quad \text{the number of constraints is } n \\ y_1 \leq 0 \quad \text{the number of constraints is } m \\ -\infty < y_2 < +\infty \quad \text{the number of constraints is } s \\ y_3 \geq 0 \quad \text{the number of constraints is } n \\ y_4 \leq 0 \quad \text{the number of constraints is } n \end{array} \right. \quad (2.49)
 \end{aligned}$$

where  $y_1 = (y_i)_{m \times 1}, y_2 = (y_i)_{s \times 1}, y_3 = (y_i)_{n \times 1}, y_4 = (y_i)_{n \times 1}$  and  $y = (y_1^T, y_2^T, y_3^T, y_4^T)^T$ .

It is obvious from model (2.49) that transformed “max” model is not a linear programming, but a nonlinear programming with inner product of different variables (wTy2). This special optimization model is also called bilinear programming. Certainly, the interior point method of nonlinear programming is only to obtain local optimal solution, but it is a NP difficult problem to obtain global optimal solution. For NP difficult problem, a unified and effective method to solve the problem is not proposed at present. In general, different approximate methods are used to solve the problem, and the artificial intelligence method and branch and bound method are mainly used to search. Two branch and bound methods are proposed in the paper to solve this problem.

**Method One: Spatial branch and bound method**

The first method to search global optimal solution is to use branch and bound method based linear relaxation technique. It is branched for continuous variables, so it is also called spatial branch and bound method or simplicial branch and bound method [7, 8].

Generally, the objective function of standard model is normalized as “min” form. Therefore, (2.49) is replaced by (2.50), which have same optimal solutions, and the model with “min” form is used to analyze and discuss below.

$$\min_{w, y} \quad -Z^+ = -w^T y_1 - (d^-)^T y_2 - l^T y_3 - u^T y_4 \quad (2.50)$$



In the  $k$ th bounded closet set  $R^k$ , the linear relaxation model of (2.49) is formulated as following:

$$\min_{\mathbf{w}, \mathbf{y}} \quad -Z^+ = -\mathbf{s} - (\mathbf{d}^-)^T \mathbf{y}_2 - \mathbf{l}^T \mathbf{y}_3 - \mathbf{u}^T \mathbf{y}_4 \quad (2.51)$$

$$\text{set} \quad \mathbf{A}^T \mathbf{y}_1 + \mathbf{B}^T \mathbf{y}_2 + \mathbf{y}_3 + \mathbf{y}_4 = \mathbf{c} \quad (2.52)$$

$$\mathbf{y}_2 \leq 0 \quad \mathbf{y}_3 \geq 0 \quad \mathbf{y}_4 \leq 0 \quad (2.53)$$

$$s_i^k \geq l_{y_1,i}^k w_i + l_{w,i}^k y_{1,i} - l_{w,i}^k l_{y_1,i}^k \quad (2.54)$$

$$s_i^k \geq u_{y_1,i}^k w_i + u_{w,i}^k y_{1,i} - u_{w,i}^k u_{y_1,i}^k \quad (2.55)$$

$$s_i^k \leq u_{y_1,i}^k w_i + l_{w,i}^k y_{1,i} - l_{w,i}^k u_{y_1,i}^k \quad (2.56)$$

$$s_i^k \leq l_{y_1,i}^k w_i + u_{w,i}^k y_{1,i} - u_{w,i}^k l_{y_1,i}^k \quad (2.57)$$

$$R^k = \left\{ \mathbf{w}^k, \mathbf{y}_1^k \in \mathbb{R}^{N_g \times 1} \mid l_w^k \leq \mathbf{w}^k \leq u_w^k, l_{y_1}^k \leq \mathbf{y}_1^k \leq u_{y_1}^k \right\} \quad (2.58)$$

The flow chart of spatial branch and bound method is given in Table 2.2. To express conveniently,  $\Psi_{\text{lb}}(\Omega)$  denotes the lower bound solution of region  $\Omega$  obtained by relaxation model (2.43), and  $\Psi_{\text{ub}}(\Omega)$  denotes the upper bound solution obtained by interior point method.

**Method Two: Integral branch and bound method**

**Theorem 2.3 [9]:** *If the polyhedral feasible region  $\mathbf{W}$  and  $\mathbf{Y}$  are bounded and separable, the optimal solution  $(\mathbf{w}^*, \mathbf{y}^*)$  of bi-linear programming with  $\mathbf{w} \in \mathbf{W}$  and  $\mathbf{y} \in \mathbf{Y}$  is satisfied with  $\mathbf{w}^* \in V(\mathbf{W})$  and  $\mathbf{y}^* \in V(\mathbf{Y})$ , where  $V(\mathbf{W})$  and  $V(\mathbf{Y})$  denotes the vertexes (poles) of polyhedral feasible region  $\mathbf{W}$  and  $\mathbf{Y}$ , respectively, then we have*

**Table 2.2** The flow chart of spatial branch and bound method

Spatial Branch and Bound Method	
Step 1	$k \leftarrow 0$ , initial region $\Omega_0$ and precision $\varepsilon$ is given;
Step 2	$L_0 = L_b(\Omega_0)$ , $U_0 = \Psi_{\text{ub}}(\Omega_0)$ , $\text{List} = \{\Omega_0\}$ ;
Step 3	<b>while</b> $U_k - L_k > \varepsilon$
Step 4	$k = \arg_j(\Psi_{\text{ub}}(\Omega_j) = L_k), \forall \Omega_j \in \text{List}$ ;
Step 5	Halve region $\Omega_k$ , and form two new region $\Omega_l$ and $\Omega_{ll}$ ;
Step 6	Remove $\Omega_k$ of List, and add $\Omega_l$ and $\Omega_{ll}$ ;
Step 7	$L_{k+1} \leftarrow \min \Psi_{\text{lb}}(\Omega_j), \forall \Omega_j \in \text{List}$
Step 8	$U_{k+1} \leftarrow \min \Psi_{\text{ub}}(\Omega_j), \forall \Omega_j \in \text{List}$
Step 9	$k \leftarrow k+1$ ;
Step 10	<b>end</b>

$$\max_{\mathbf{w} \in \mathbf{W}} f(\mathbf{w}, \mathbf{y}^*) = f(\mathbf{w}^*, \mathbf{y}^*) = \max_{\mathbf{y} \in \mathbf{Y}} f(\mathbf{w}^*, \mathbf{y}) \quad (2.59)$$

*Proof* If the value of fixed vector  $\mathbf{y}$  is  $\mathbf{y}^*$ , the bilinear programming model is transformed into a linear programming with variable  $\mathbf{w}$  and its optimal solution  $\mathbf{w}^*$  must be at the poles of the feasible region of  $\mathbf{W}$ . Similarly, if the value of fixed vector  $\mathbf{w}$  is  $\mathbf{w}^*$ , then the bilinear programming model is transformed into a linear programming with variable  $\mathbf{y}$  and its optimal solution  $\mathbf{y}^*$  must be at the poles of the feasible region of  $\mathbf{Y}$ . It notes that the optimal solution  $(\mathbf{w}^*, \mathbf{y}^*)$  of the bilinear programming model is satisfied with  $\mathbf{w}^* \in V(\mathbf{W})$  and  $\mathbf{y}^* \in V(\mathbf{Y})$ , where  $V(\mathbf{W})$  and  $V(\mathbf{Y})$  represents the vertexes (poles) of polyhedral feasible region  $\mathbf{W}$  and  $\mathbf{Y}$ , respectively.

Based on this theorem, for the bilinear programming model of (2.49). The feasible region of uncertainties is a boxed polyhedron feasible region, so the optimal solution must be achieved at the interval upper or lower bounds of the uncertainties. Therefore, “or” constraint can be used to express. Furthermore, with a dummy integer variable  $z$ , “or” constraint can be transformed into a mixed integer linear programming model, such that

$$\begin{aligned} w_i = b_i^- \quad \text{or} \quad w_i = b_i^+ \\ \text{for } \forall i = 1, \dots, s \end{aligned} \Leftrightarrow w_i = \underline{b_i} + z_i(\overline{b_i} - \underline{b_i}), z_i \in \{0, 1\}, \quad \forall i = 1, \dots, s \quad (2.60)$$

Furthermore, the pessimistic model of (2.49) can be transformed into

$$\begin{aligned} \max_{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{w}} \quad & Z^+ = (\mathbf{d}^-)^T \mathbf{y}_1 + \mathbf{w}^T \mathbf{y}_2 + \mathbf{L}^T \mathbf{y}_3 + \mathbf{U}^T \mathbf{y}_4 \\ \text{(Pessimistic value)} \quad & \text{s.t.} \quad \mathbf{E}^T \mathbf{y}_1 + \mathbf{A}^T \mathbf{y}_2 + \mathbf{y}_3 + \mathbf{y}_4 = \mathbf{c} \\ & \mathbf{y}_1 \leq 0 \quad -\infty < \mathbf{y}_2 < +\infty \quad \mathbf{y}_3 \geq 0 \quad \mathbf{y}_4 \leq 0 \\ & w_i = b_i^- + z_i(b_i^+ - b_i^-), \quad z_i \in \{0, 1\} \end{aligned} \quad (2.61)$$

We can know that, the model (2.61) is a mixed integer bilinear programming. Fortunately, based on big M theory, the product  $zy_2$  of integer variable and continuous variable can be introduced some dummy variables, and the model is simplified as a simple dummy integer variable as following:

$$\begin{aligned} \text{(Pessimistic value)} \quad & \max_{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \mathbf{w}} \quad Z^+ = (\mathbf{d}^-)^T \mathbf{y}_1 + \sum_{i=1}^s (y_{2,i} b_i^- + x_i(b_i^+ - b_i^-)) + \mathbf{L}^T \mathbf{y}_3 + \mathbf{U}^T \mathbf{y}_4 \\ & \text{s.t.} \quad \mathbf{E}^T \mathbf{y}_1 + \mathbf{A}^T \mathbf{y}_2 + \mathbf{y}_3 + \mathbf{y}_4 = \mathbf{c} \\ & \mathbf{y}_1 \leq 0 \quad -\infty < \mathbf{y}_2 < +\infty \quad \mathbf{y}_3 \geq 0 \quad \mathbf{y}_4 \leq 0, \quad z_i \in \{0, 1\} \\ & y_{2,i} - M(1 - z_i) \leq x_i \leq y_{2,i} + M(1 - z_i), \quad -Mz_i \leq x_i \leq Mz_i \end{aligned} \quad (2.62)$$



where  $M$  is a sufficient large number. We can know that if  $z_i = 0$ ,  $x_i$  is 0, and the value of  $w_i y_{2i}$  is  $y_{2,i} b_i^-$ ; if  $z_i = 1$ ,  $x_i$  is  $y_{2,i}$ , and the value of  $w_i y_{2i}$  is  $y_{2,i} b_i^+$ .

The general solution of mixed integer programming is to relax integer variables into continuous variables, branch and bound next, and stop searching when the upper bound and lower bound are almost equal. Because the optimal solution of (2.49) is solution of the dual problem of model (2.40)–(2.44), the optimal solution of original model (2.40)–(2.44) can be solved by complementary relaxation condition. Another method is to obtain the optimal solution  $w^p$  of uncertainties based on (2.62), then fix and bring it into model (2.49). Due to uncertainties defined, the model can be transformed into:

(Pessimistic value)

$$\min_x \quad Z^+ = c^T x \quad (2.63)$$

$$\text{s.t.} \quad Ax = w^p \quad (2.64)$$

$$Bx \leq d^- \quad (2.65)$$

$$l \leq x \leq u \quad (2.66)$$

It is necessary to know that optimistic value and pessimistic value of interval optimization model are obtained according to (2.45)–(2.48) and (2.63)–(2.66), respectively, and the optimal interval of objective function is  $[Z', Z'']$ . Interval optimization mainly focus on the interval upper and lower bound information of objective function, so the optimal solution  $x^{opt}$  and  $x^{opt}$  of interval independent variables reflect the objective change from the optimistic value to the pessimistic value, which can not define interval range of optimal independent variables strictly. To obtain interval of optimal independent variables, a bi-level optimization model can be formulated as:

$$\begin{aligned} x_i^\pm &= \max_{x,w,d} / \min_{x,w,d} x_i \\ \text{s.t.} \quad & x_i^{\min} \leq x_i \leq x_i^{\max} \\ & w^- \leq w \leq w^+ \\ & d^- \leq d \leq d^+, \quad i = 1, \dots, n \\ & x_i = \arg \min_x c^T x \\ \text{s.t.} \quad & Ax = w \\ & Bx \leq d \end{aligned} \quad (2.67)$$

The combinatorial nature of bi-level programming can be observed by studying the single-level reformulation obtained by replacing the inner level problem with its KKT (Karush-Kuhn-Tucker) condition, which is a necessary and sufficient optimality conditions when the inner optimization model is actually a convex problem.

However, KKT condition is a set of nonlinear and non-convex equations; therefore, though KKT condition is used, it is hard to guarantee the global optimality of programming. Then we consider to transform the KKT condition into an expression of mixed integer, and by introducing some dummy variables, nonlinear and non-convex equations can be replaced by non-convex equations of mixed integer linear programming, which is easier to solve. Furthermore, combined big M theory, the general complementarity slackness constraints can be exactly linearized with dummy integer variables  $z$ , such that (2.68) is equivalently transformed as (2.69).

$$\begin{aligned}
 & \mathbf{Ax} = \mathbf{w} \\
 & \mathbf{Bx} \leq \mathbf{d} \\
 & \mathbf{c} + \mathbf{A}^T \boldsymbol{\mu} + \mathbf{B}^T \boldsymbol{\lambda} = \mathbf{0} \\
 & (\mathbf{Bx} - \mathbf{d}) \cdot * \boldsymbol{\lambda} = \mathbf{0} \\
 & \boldsymbol{\lambda} \geq \mathbf{0}
 \end{aligned} \tag{2.68}$$

$$\begin{aligned}
 & \mathbf{Ax} = \mathbf{w} \\
 & \mathbf{c} + \mathbf{A}^T \boldsymbol{\mu} + \mathbf{B}^T \boldsymbol{\lambda} = \mathbf{0} \\
 & -M(1 - z) \leq (\mathbf{Bx} - \mathbf{d}) \leq \mathbf{0} \\
 & Mz \geq \boldsymbol{\lambda} \geq \mathbf{0}, z \in \{0, 1\}
 \end{aligned} \tag{2.69}$$

Furthermore, the model (2.67) is transformed into

$$\begin{aligned}
 x_i^\pm &= \max_x / \min_x \\
 & \text{s.t.} \\
 & x_i^{\min} \leq x_i \leq x_i^{\max} \\
 & \mathbf{w}^- \leq \mathbf{w} \leq \mathbf{w}^+ \\
 & \mathbf{d}^- \leq \mathbf{d} \leq \mathbf{d}^+ \\
 & \mathbf{Ax} = \mathbf{w} \\
 & \mathbf{c} + \mathbf{A}^T \boldsymbol{\mu} + \mathbf{B}^T \boldsymbol{\lambda} = \mathbf{0} \\
 & -M(1 - z) \leq \mathbf{Bx} - \mathbf{d} \leq \mathbf{0} \\
 & Mz \geq \boldsymbol{\lambda} \geq \mathbf{0}, \quad z \in \{0, 1\}
 \end{aligned}, \quad i = 1, \dots, n \tag{2.70}$$

Moreover, the mixed integer linear programming can be solved by branch and bound method to obtain upper and lower bound of optimal solution as well as its interval  $[x_i^-, x_i^+]$ .

However, when conducting interval optimization, it is necessary to consider such a problem that whether the objective function and optimal solution are continuous functions of uncertainties or not. In other words, when the uncertainties are changed in the range of interval, the objective function and the optimal solution are also changed in a certain range, but whether they are filled with the entire range of interval. Next, we will prove that if the objective function and optimal solution are continuous functions of uncertainties, then they can be filled with the entire range of interval.

Mathematically, linear programming (2.28)–(2.30) can be formulated as

$$\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{q}^T \mathbf{x} + r \quad (2.71)$$

$$s.t. \quad \mathbf{B}_{eq} \mathbf{x} = \mathbf{g}_{eq} + \mathbf{F}_{eq} \boldsymbol{\theta} \quad (2.72)$$

$$\mathbf{B}_{ieq} \mathbf{x} \leq \mathbf{g}_{ieq} + \mathbf{F}_{ieq} \boldsymbol{\theta} \quad (2.73)$$

$$\boldsymbol{\theta} \in [\boldsymbol{\theta}^-, \boldsymbol{\theta}^+] \quad (2.74)$$

where  $\mathbf{x}$  is an  $n \times 1$  vector;  $\boldsymbol{\theta}$  is an  $t \times 1$  parametric vector with its upper bound  $\boldsymbol{\theta}^+$  and lower bound  $\boldsymbol{\theta}^-$ , respectively;  $\mathbf{Q}$  is an  $n \times n$  positive definite matrix;  $\mathbf{q}$  is an  $n \times 1$  vector;  $\mathbf{B}_{eq}$  is an  $m \times n$  coefficient matrix;  $m$  is the number of equality constraints;  $\mathbf{g}_{eq}$  is an  $m \times 1$  constant vector;  $\mathbf{F}_{eq}$  is an  $m \times t$  coefficient matrix;  $\mathbf{B}_{ieq}$  is an  $l \times n$  coefficient matrix;  $l$  is the number of inequality constraints;  $\mathbf{g}_{ieq}$  is an  $l \times 1$  constant vector;  $\mathbf{F}_{ieq}$  is an  $l \times t$  coefficient matrix.

The KKT conditions of model (2.71)–(2.74) are formulated as

$$\mathbf{B}_{eq}^T \boldsymbol{\pi} + \mathbf{B}_{ieq}^T \boldsymbol{\mu} + \mathbf{q} = \mathbf{0} \quad (2.75)$$

$$\mathbf{B}_{eq} \mathbf{x} - \mathbf{g}_{eq} - \mathbf{F}_{eq} \boldsymbol{\theta} = \mathbf{0} \quad (2.76)$$

$$\mathbf{B}_{ieq} \mathbf{x} - \mathbf{g}_{ieq} - \mathbf{F}_{ieq} \boldsymbol{\theta} \leq \mathbf{0} \quad (2.77)$$

$$\mu^i \left( B_{ieq}^i \mathbf{x} - g^i + F^i \boldsymbol{\theta} \right) = 0 \quad (2.78)$$

$$\boldsymbol{\pi} \geq \mathbf{0} \quad (2.79)$$

**Definition:** For each constraint of (2.75), we define the  $i$ th constraint to be an active constraint with  $B_{ieq}^i \mathbf{x} - g^i + F^i \boldsymbol{\theta} = 0$ , and inactive constraint with  $B_{ieq}^i \mathbf{x} - g^i + F^i \boldsymbol{\theta} < 0$ . Furthermore, the optimal active set  $\mathcal{A}$  is defined as the set of indices of active constraints at the optimum, such that  $\mathcal{A} = \left\{ i \mid B_{ieq}^i \mathbf{x} - g^i + F^i \boldsymbol{\theta} = 0 \right\}$  and the optimal inactive set  $\mathcal{I}$  is defined as the set of indices of active constraints at the optimum, such that  $\mathcal{I} = \left\{ i \mid B_{ieq}^i \mathbf{x} - g^i + F^i \boldsymbol{\theta} < 0 \right\}$ .

For the active set  $\mathcal{A}$ , we have (2.80) and the corresponding Lagrange multipliers are  $\boldsymbol{\mu}^{\mathcal{A}} \geq \mathbf{0}$ .

$$\mathbf{B}_{ieq}^{\mathcal{A}} \mathbf{x} - \mathbf{g}_{ieq}^{\mathcal{A}} - \mathbf{F}_{ieq}^{\mathcal{A}} \boldsymbol{\theta} = \mathbf{0} \quad (2.80)$$

For the inactive set  $\mathcal{I}$ , we have (2.81) and the corresponding Lagrange multipliers are  $\boldsymbol{\mu}^{\mathcal{I}} = \mathbf{0}$ .

$$\mathbf{B}_{ieq}^T \mathbf{x} - \mathbf{g}_{ieq}^T - \mathbf{F}_{ieq}^T \boldsymbol{\theta} < \mathbf{0} \quad (2.81)$$

Furthermore, we assume the original linear optimization model is feasible and the constraints are satisfied, such that  $\begin{pmatrix} \mathbf{B}_{eq} \\ \mathbf{B}_{ieq}^A \end{pmatrix}$  has full row rank. According to (2.76) and (2.80), we have

$$\begin{pmatrix} \mathbf{B}_{eq} \\ \mathbf{B}_{ieq}^A \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{g}_{eq} \\ \mathbf{g}_{ieq}^A \end{pmatrix} + \begin{pmatrix} \mathbf{F}_{eq} \\ \mathbf{F}_{ieq}^A \end{pmatrix} \boldsymbol{\theta} \quad (2.82)$$

Then  $\mathbf{x}$  yields

$$\mathbf{x} = \begin{pmatrix} \mathbf{B}_{eq} \\ \mathbf{B}_{ieq}^A \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{g}_{eq} \\ \mathbf{g}_{ieq}^A \end{pmatrix} + \begin{pmatrix} \mathbf{B}_{eq} \\ \mathbf{B}_{ieq}^A \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{F}_{eq} \\ \mathbf{F}_{ieq}^A \end{pmatrix} \boldsymbol{\theta} \quad (2.83)$$

From (2.83), we can find that the active set  $\mathcal{A}$  changes with parameter  $\boldsymbol{\theta}$ . Furthermore, the optimal solution  $\mathbf{x}$  is a piecewise linear function. According to KKT conditions, because the model is a convex programming, the optimal solution of (2.75)–(2.79) obtained by KKT conditions is the global optimal solution of linear programming (2.71)–(2.74) sufficiently and necessarily, and the optimal solution is unique. Therefore, the optimal solution  $\mathbf{x}$  is a piecewise linear continuous function, which further illustrates the optimal solution is the continuous function of uncertainties and they must be able to fill the entire interval range. Take the optimal solution  $\mathbf{x}$  into the objective function of (2.71), we can know the objective function is a piecewise linear continuous function about  $\boldsymbol{\theta}$ , which can also fill the entire interval range.

### 2.3.2 Traditional Interval Robust Optimization Method

Firstly, a traditional robust optimization model is formulated as

$$\min_{\mathbf{x}} \quad \mathbf{a}^T \mathbf{x} \quad (2.84)$$

$$s.t. \quad \mathbf{b} \leq \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \leq \mathbf{d} \quad (2.85)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.86)$$

$$\forall \mathbf{u} \in \Omega_{\mathbf{u}} \quad (2.87)$$

where  $\Omega_{\mathbf{u}}$  denotes the set of uncertainties, if it is an interval number, such as  $\Omega_{\mathbf{u}} = \{\mathbf{u} | \mathbf{u}^- \leq \mathbf{u} \leq \mathbf{u}^+\}$ , the model is called a traditional interval robust

optimization model. The objective of model is to search an optimal  $\mathbf{x}$ , and satisfy constraint (2.85) with any change of uncertainty  $\mathbf{u}$ . The model (2.84)–(2.87) can be transformed into a bi-level programming model as following:

$$\min_{\mathbf{x}} \max_{\mathbf{u} \in \Omega_{\mathbf{u}}} \mathbf{a}^T \mathbf{x} \quad (2.88)$$

$$s.t. \quad \mathbf{b} \leq \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \leq \mathbf{d} \quad (2.89)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.90)$$

Because the objective function is not related to  $\mathbf{u}$ , then (2.88)–(2.90) is equivalent to:

$$\min_{\mathbf{x}} \mathbf{a}^T \mathbf{x} \quad (2.91)$$

$$s.t. \quad \mathbf{A}\mathbf{x} + \max_{\mathbf{u} \in \Omega_{\mathbf{u}}} (\mathbf{B}\mathbf{u}) \leq \mathbf{d} \quad (2.92)$$

$$\mathbf{A}\mathbf{x} + \min_{\mathbf{u} \in \Omega_{\mathbf{u}}} (\mathbf{B}\mathbf{u}) \geq \mathbf{b} \quad (2.93)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.94)$$

Considering the interval form of uncertainties in model (2.92) and (2.93),  $\max_{\mathbf{u} \in \Omega_{\mathbf{u}}} (\mathbf{B}\mathbf{u}) = \mathbf{B}^+ \mathbf{u}^+ + \mathbf{B}^- \mathbf{u}^-$ ,  $\min_{\mathbf{u} \in \Omega_{\mathbf{u}}} (\mathbf{B}\mathbf{u}) = \mathbf{B}^+ \mathbf{u}^- + \mathbf{B}^- \mathbf{u}^+$ , where  $\mathbf{B}^+ = \max(\mathbf{B}, \mathbf{0})$ , such that  $\mathbf{B}^+$  is the positive elements of  $\mathbf{B}$ , and  $\mathbf{B}^- = \min(\mathbf{B}, \mathbf{0})$ , such that  $\mathbf{B}^-$  is the negative elements of  $\mathbf{B}$ . Furthermore, (2.91)–(2.94) is translated into a bi-level optimization model:

$$\min_{\mathbf{x}} \mathbf{a}^T \mathbf{x} \quad (2.95)$$

$$s.t. \quad \mathbf{A}\mathbf{x} + \max(\mathbf{B}, \mathbf{0})\mathbf{u}^+ + \min(\mathbf{B}, \mathbf{0})\mathbf{u}^- \leq \mathbf{d} \quad (2.96)$$

$$\mathbf{A}\mathbf{x} + \max(\mathbf{B}, \mathbf{0})\mathbf{u}^- + \min(\mathbf{B}, \mathbf{0})\mathbf{u}^+ \geq \mathbf{b} \quad (2.97)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.98)$$

It notes that an important characteristic of the model is no equality constraints, because traditional robust optimization model requires that constraints can not be violated with any change of uncertainty  $\mathbf{u}$ . if equality constraints are existed, they must be violated and robust optimal solution can be obtained. On the other hand, from model (2.88)–(2.90), we find that the transformation result is the same as the pessimistic value of interval optimization without equality constraints. It illustrates that although pessimism value has some sacrifices of objective, it can ensure all constraints are satisfied with any change of uncertainties.

An simple example:

$$\min_{x_1+x_2} 0.1x_1 + 0.2x_2 \quad s.t. \begin{cases} 0 \leq 2x_1 + x_2 + u \leq 6 \\ 0 \leq x_1 \leq 3 \\ 0 \leq x_2 \leq 3 \end{cases}, \quad \forall u \in [-1, +1]$$

The objective of the optimization model is to obtain an optimal solution  $(x_1^*, x_2^*)$ , which can ensure all constraints satisfied for any given  $u \in [-1, +1]$ . The model is equivalent to:

$$\begin{aligned} \min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^* \quad s.t. \begin{cases} \max_{-1 \leq u \leq +1} (-u) \leq 2x_1^* + x_2^* \leq \min_{-1 \leq u \leq +1} (6 - u) \\ 0 \leq x_1^* \leq 3 \\ 0 \leq x_2^* \leq 3 \end{cases}, \text{ furthermore, simplified as} \\ \min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^* \quad s.t. \begin{cases} 1 \leq 2x_1^* + x_2^* \leq 5 \\ 0 \leq x_1^* \leq 3 \\ 0 \leq x_2^* \leq 3. \end{cases} \end{aligned}$$

### 2.3.3 Adaptive Interval Robust Optimization Method

In traditional interval robust optimization model, equality constraints can not be considered because the model need to obtain an optimal solution  $\mathbf{x}$  which can satisfy all constraints with any change of uncertainty  $\mathbf{u}$ ; therefore, equality constraints are not always satisfied. On the other hand, traditional interval robust optimization requires that optimal variable  $\mathbf{x}$  is not related to uncertainty  $\mathbf{u}$ . If equality constraints are existed, we can obtain an optimal variable  $\mathbf{x}$  irrelevant to  $\mathbf{u}$ .

However, considering that the independent variable  $\mathbf{x}$  in the model can follow the change of uncertainty  $\mathbf{u}$ , such that  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$ , then the optimal  $\mathbf{x}$  can be solved only if the adjusted variable  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$  can guarantee that all constraints are not violated with any given uncertainty  $\mathbf{u}$ . In such situation, after the change of uncertainty  $\mathbf{u}$ , the equality constraint can be satisfied because the independent variable  $\mathbf{x}$  is changed accordingly. Such model is called adaptive robust optimization model shown following:

$$\min_{\mathbf{x}} \mathbf{a}^T \mathbf{x} \quad (2.99)$$

$$s.t. \mathbf{A}\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u}) + \mathbf{B}\mathbf{u} = \mathbf{w} \quad (2.100)$$

$$\mathbf{b} \leq \mathbf{C}\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u}) + \mathbf{D}\mathbf{u} \leq \mathbf{d} \quad (2.101)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.102)$$

$$\forall \mathbf{u} \in \Omega_{\mathbf{u}} \quad (2.103)$$

Similarly, if the set of uncertainties is an interval form, the model is called an adaptive interval robust optimization model which has challenge to be solved. Generally, the equality constraints should be analyzed and  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$  related to uncertainty is strictly expressed, according to the physically practical problems or the structural characteristics of the model. If  $\mathbf{u}^0$  is the given value of the uncertainty, satisfied with  $\mathbf{Ax} + \mathbf{Bu}^0 = \mathbf{w}$ , then consider  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u}) = \mathbf{x} + \mathbf{L}(\mathbf{u}^0 - \mathbf{u})$ , and satisfy  $\mathbf{AL}(\mathbf{u}^0 - \mathbf{u}) = \mathbf{B}(\mathbf{u}^0 - \mathbf{u})$ . It illustrates that when the uncertainty is deviated from  $\mathbf{u}^0$ , the independent variable  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$  can be adjusted accordingly to guarantee the equality constraints, where  $\mathbf{L}$  is usually a linear function called the affine expansion of uncertainty  $\mathbf{u}$ . Nevertheless, if  $\mathbf{L}$  is given, with any change of  $\mathbf{u}$  in the  $\Omega_{\mathbf{u}}$ , we should guarantee that  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$  can always satisfy equality constraints as well as inequality constraints. Furthermore, the adaptive robust optimization model can be transformed into traditional robust optimization model as following:

$$\min_{\mathbf{x}} \quad \mathbf{a}^T \mathbf{x} \quad (2.104)$$

$$s.t. \quad \mathbf{Ax} + \mathbf{Bu}^0 = \mathbf{w} \quad (2.105)$$

$$\mathbf{b} \leq \mathbf{Cx} + (\mathbf{CL}(\mathbf{u}^0 - \mathbf{u}) + \mathbf{Du}) \leq \mathbf{d} \quad (2.106)$$

$$\mathbf{x} \in \Omega_{\mathbf{x}} \quad (2.107)$$

$$\forall \mathbf{u} \in \Omega_{\mathbf{u}} \quad (2.108)$$

We should know that the biggest challenge of adaptive robust optimization model is how to find the strict expression of  $\tilde{\mathbf{x}}(\mathbf{x}, \mathbf{u})$  about uncertainty, such that. If it can be obtained, then the problem is easily solved.

A simple example:

$$\min_{x_1, x_2} \quad 0.1x_1 + 0.2x_2 \quad s.t. \begin{cases} x_1 + x_2 + u = 4 \\ 0 \leq 2x_1 + x_2 + u \leq 6 \\ 0 \leq x_1 \leq 3 \\ 0 \leq x_2 \leq 3 \end{cases}, \quad \forall u \in [-1, +1]$$

The objective of the optimization model is also to obtain an optimal solution  $(x_1^*, x_2^*)$ , which can ensure all constraints satisfied for any given  $u \in [-1, +1]$ . However, due to equality constraints, constraints are not always satisfied; therefore, the traditional robust optimization model can not be obtained. Considering the change of uncertainties,  $x_1$  and  $x_2$  is changed accordingly, such that  $\begin{cases} x_1 = x_1^* + L_1(u^0 - u) \\ x_2 = x_2^* + L_2(u^0 - u) \end{cases}$ .

Considering the linear expansion of  $L_1$  and  $L_2$ , such as  $\begin{cases} x_1 = x_1^* + 0.2(u^0 - u) \\ x_2 = x_2^* + 0.8(u^0 - u) \end{cases}$ , then we find  $x_1 + x_2 + u = x_1^* + x_2^* + u^0$ . Therefore, no matter how uncertainty  $u$  changes, as long as  $x_1^* + x_2^* + u^0$  is satisfied, equality constraints will be always satisfied with the linear adjustment of  $x_1$  and  $x_2$ . To guarantee inequality constraints

satisfied, we have  $\min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^* \quad s.t. \quad \begin{cases} x_1^* + x_2^* + u^0 = 4 \\ 0 \leq 2(x_1^* + 0.2(u^0 - u)) + x_2^* + 0.8(u^0 - u) + u \leq 6 \\ 0 \leq x_1^* \leq 3 \\ 0 \leq x_2^* \leq 3 \end{cases}$ ,

$\forall u \in [-1, +1]$  it can be further simplified as  $\min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^*$   
 $s.t. \quad \begin{cases} x_1^* + x_2^* + u^0 = 4 \\ 0 \leq 2x_1^* + x_2^* + 1.2u^0 - 0.2u \leq 6, \forall u \in [-1, +1]. \end{cases}$  We can know that the model is the

same as traditional robust optimization and uncertainty is only existed in the inequality constrain, so it can be transformed into  $\min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^*$

$s.t. \quad \begin{cases} x_1^* + x_2^* + u^0 = 4 \\ 0.2 \leq 2x_1^* + x_2^* + 1.2u^0 \leq 5.8 \\ 0 \leq x_1^* \leq 3 \\ 0 \leq x_2^* \leq 3 \end{cases}$ . Generally,  $u^0$  is considered as the desired value

of  $u$ , for example  $u^0 = 0$ , and the model is formulated as  $\min_{x_1^*, x_2^*} 0.1x_1^* + 0.2x_2^*$

$s.t. \quad \begin{cases} x_1^* + x_2^* = 4 \\ 0.2 \leq 2x_1^* + x_2^* \leq 5.8 \\ 0 \leq x_1^* \leq 3 \\ 0 \leq x_2^* \leq 3 \end{cases}$ . It notes that  $L_1$ ,  $L_2$  and  $u^0$  of the model should be

determined based on the actual situation, and the different choices of  $L_1$ ,  $L_2$  and  $u^0$  lead to different optimal results.

### 2.3.4 Two-Stage Interval Robust Optimization Method

Generally, a two-stage robust optimization model can be formulated as

$$\min_{y \in \Omega_y} c^T y + \max_{u \in \Omega_u} \min_{x \in \Omega_x} a^T x \quad (2.109)$$

where  $\Omega_y$  and  $\Omega_u$  are both bounded convex sets and  $\Omega_x = \{x | Ax \geq d - By - Cu\}$ . If uncertainty set is an interval number, then the model is call two-stage interval robust optimization model.

The objective of two-stage robust optimization is to obtain optimal variable  $y$  in the first stage and optimal variable  $x^*$  in the second stage. Actually,  $x^*$  is a function of uncertainty  $u$ , such that  $x^*(u)$ , so the optimal variable and objective function are changed accordingly with the uncertainties. Robust optimization is to choose  $u$  which can obtain the best  $y$  when  $x^*(u)$  achieves the worst case. If the uncertainty



set is modeled with intervals, the two-stage robust optimization model based intervals is formed.

The first-stage variable of the two-stage interval robust optimization model is generally discrete variable, and the second-stage variable is continuous variable. Therefore, the two-stage interval robust optimization model is a large-scale combinatorial optimization problem which can be solved by decomposition algorithms, where Benders decomposition method is an effective method and is widely used [10–13]. For Benders decomposition method, the original problem is decomposed into master problem and subproblem, by solving each subproblem, generating Benders cut and adding it to the master problem, and the master problem is solved to get a better objective function, through several iterations, then the optimal solution is obtained. In the flow chart of Benders, the most important step is to form Benders cut, and different cut generation methods have different computational complexity. Two kinds of Benders cut generation methods are presented.

The inner level of “max-min” is the same as pessimistic model of interval optimization, so the corresponding method can be used, which is to solve the duality of “min” model in the “max-min”. The dual model is

$$(SP) \quad \Theta(y) = \max_{u, \pi} \pi^T (d - By - Cu) \quad (2.110)$$

$$s.t. \quad A^T \pi \leq a^T \quad (2.111)$$

$$u \in \Omega_u, \quad \pi \geq 0 \quad (2.112)$$

It shows that (SP) is the subproblem of Benders which is an iterative algorithm. In the  $k$ th iteration, for a given  $y_k^*$ , the subproblem  $\Theta(y_k^*)$  can be solve to obtain the optimal solution  $(u_k^*, \pi_k^*)$ . Then two kinds of cuts are designed for iterations.

#### A. Benders decomposition method based on dual cut

By solving subproblem SP, the dual information  $\pi_k^*$  of inner problem is obtained. Based on weak duality theorem, the dual problem of the linear programming model must be a lower bound of the original problem, so the dual cut is

$$\eta \geq (\pi_k^*)^T (d - By - Cu_k^*) \quad (2.113)$$

Add the dual cut (2.113) to Benders main problem, we have

$$(MP1) \quad \min_{y, \eta} \quad c^T y + \eta \quad (2.114)$$

$$s.t. \quad \eta \geq (\pi_l^*)^T (d - By - Cu_l^*), \quad l = 1, 2, \dots, k \quad (2.115)$$

$$\mathbf{y} \in \Omega_{\mathbf{y}} \quad (2.116)$$

We can see the main problem is a simple mixed integer programming model, which can be solved by branch and bound method, and  $(\mathbf{y}_{k+1}^*, \eta_{k+1}^*)$  denotes to the optimal solution. The objective functions  $\mathbf{c}^T \mathbf{y}_k^* + \Theta(\mathbf{y}_k^*)$  and  $\mathbf{c}^T \mathbf{y}_{k+1}^* + \eta_{k+1}^*$  of MP1 give the upper and lower bound of original optimization model, respectively. By continuous iterations, subproblem increases optimal cuts gradually and the upper bound and lower bound are finally reaching consistently, so the algorithm converges to the global optimal solution.

#### B. Benders decomposition method based on original cut

The inner-level “max-min” model of original three-level optimization problem is transformed into a max bilinear programming with duality. The specific structure of bilinear programming discussed in the interval optimization is that the bilinear variable  $(\mathbf{u}, \boldsymbol{\pi})$  can be separable, which illustrates that the global optimal solution must be achieved at the vertices (or poles) of the defined feasible region  $\mathbf{u}$  and  $\boldsymbol{\pi}$ . The optimal solution  $(\mathbf{u}_k^*, \boldsymbol{\pi}_k^*)$  obtained in  $k$ th the iteration is only an optimal solution with the given  $\mathbf{y}_k$ , so an original cut is designed as

$$\mathbf{A}\mathbf{x}^l \geq \mathbf{d} - \mathbf{B}\mathbf{y} - \mathbf{C}\mathbf{u}_l^*, \quad \eta \geq \mathbf{a}^T \mathbf{x}^l \quad (2.117)$$

Add the original cut (2.117) to Benders main problem, we have

$$(\text{MP2}) \quad \min_{\mathbf{y}, \eta, \mathbf{x}^l} \quad \mathbf{c}^T \mathbf{y} + \eta \quad (2.118)$$

$$s.t. \quad \eta \geq \mathbf{a}^T \mathbf{x}^l, \quad \mathbf{A}\mathbf{x}^l + \mathbf{B}\mathbf{y} \geq \mathbf{d} - \mathbf{C}\mathbf{u}_l^*, \quad l = 1, 2, \dots, k \quad (2.119)$$

$$\mathbf{y} \in \Omega_{\mathbf{y}} \quad (2.120)$$

We can see the main problem is a simple mixed integer programming model, which can be solved by branch and bound method, and denotes to the optimal solution. Similarly, the objective functions  $\mathbf{c}^T \mathbf{y}_k^* + \Theta(\mathbf{y}_k^*)$  and  $\mathbf{c}^T \mathbf{y}_{k+1}^* + \eta_{k+1}^*$  of MP2 give the upper and lower bound of original optimization model, respectively. By continuous iterations, subproblem increases optimal cuts gradually and the upper bound and lower bound are finally reaching consistently, so the algorithm converges to the global optimal solution. The difference between MP1 is when increasing the original cuts, the extra variables will be added. Therefore, with the increase of iterations, the number of variables in the model (MP2) is more and more, and the scale of the model will be increased, but the total number of iterations will be greatly reduced.

**Theorem 2.4** [14] *if  $p$  is the pole number of uncertainty set  $\Omega_u$ ,  $q$  is the pole number of simplex  $\{A^T\pi \leq a^T, \pi \geq 0\}$ , then the iteration number of Benders decomposition method based on dual cut is  $\mathcal{O}(pq)$ ; the iteration number of Benders decomposition method based on original cut is  $\mathcal{O}(p)$ .*

From 2.4, it is obvious that by using original cut, the computation complexity is greatly reduced, and the solving efficiency is  $Q$  times higher. Therefore, the method based on original cut is mainly used in this paper. Of course, in the actual calculation, we can also add the original cut and the dual cut into the model, while preserving the characteristics of the original cut and the dual cut. Next, Benders decomposition method based on original cut is refined, given convergence precision  $\varepsilon$ , and the flow chart is shown following:

- (i) Set  $LB = -\infty$ ,  $UB = +\infty$ ,  $k = 0$ ;
- (ii) Solve main problem (MP2), and obtain the optimal solution  $(y_{k+1}^*, \eta_{k+1}^*, x_1^*, \dots, x_k^*)$  and lower bound of model  $LB = \eta^*$ ;
- (iii) Fix  $y^*$  to solve subproblem (SP), if subproblem is feasible, obtain optimal solution  $(u_k^*, \pi_k^*)$  and objective value  $\Theta(y^*)$ ; if not,  $\Theta(y^*) = +\infty$ . Furthermore, update upper bound of model with  $UB = \min\{UB, \Theta(y^*)\}$ ;
- (iv) If  $(UB-LB) \leq \varepsilon$ , return  $y^*$  and stop iterations; otherwise  $k = k+1$ , fix  $u_k^*$ , design original cut:
  - (a) If the subproblem of (iii) is feasible, generate new variable  $x^l$ , and add cut shown in (2.121) to the main problem (MP);

$$Ax^l \geq d - By - Cu_l^*, \quad \eta \geq a^T x^l \quad (2.121)$$

- (b) If the subproblem of (iii) is infeasible, generate new variable  $x^l$  and add cut shown in (2.122) to the main problem (MP);

$$Ax^l \geq d - By - Cu_l^* \quad (2.122)$$

- (v) Go to (ii).

According to the flow chart of Benders decomposition method based on original cut, in the fourth step, if the subproblem is feasible, the added original cut is called optimal cut from the mathematical structure, because it is obtained with the information of optimal solution; but if the subproblem is infeasible, the original cut is changed and called feasible cut which is only to guarantee the feasibility of problem.

Another important problem is how to solve the inner-level “maxmin” optimization model. It has been illustrated anteriorly that if the uncertainty set is an interval form, the “maxmin” optimization is the same as the pessimistic model of interval optimization. Furthermore, it is proved that optimal solution must be obtained when uncertainties reach the bounds, so it can be transformed into a mixed linear programming model to solve. However, if the robust cost is considered, the

structure of uncertainty set is changed, and optimal solution is not achieved at the bound (but achieved at the vertices of the simplex). Therefore, new methods need to solve.

Assume the dual variable of second stage is  $\pi$ . Because the “maxmin” can be regarded as a bilinear programming where inner-level model is a linear programming, if the inner-level model is feasible, then KKT condition is the sufficient and necessary condition of optimal solution and we have optimization model as following:

$$\max_{x,u,\pi} 0 \quad (2.123)$$

$$s.t. \quad A^T \pi \leq a^T \quad (2.124)$$

$$Ax \geq d - By - Cu \quad (2.125)$$

$$u \in \Omega_u, \quad \pi \geq 0 \quad (2.126)$$

$$\pi_i(Ax - d + By + Cu)_i = 0, \quad \forall i \quad (2.127)$$

$$x_j(a^T - A^T \pi)_j = 0, \quad \forall j \quad (2.128)$$

$$u \in \Omega_u, \quad \pi \geq 0 \quad (2.129)$$

where constraints (2.127) and (2.128) are complementary, which can be transformed into mixed integer programming to solve based on big M theory, and we have:

$$\max_{x,u,\pi,v,w} 0 \quad (2.130)$$

$$s.t. \quad A^T \pi \leq a^T \quad (2.131)$$

$$Ax \geq d - By - Cu \quad (2.132)$$

$$u \in \Omega_u, \quad \pi \geq 0 \quad (2.133)$$

$$\pi_i \leq Mw_i, \quad (Ax - d + By + Cu)_i \leq M(1 - w_i), \quad \forall i \quad (2.134)$$

$$0 \leq x_j \leq Mv_j, \quad (a^T - A^T \pi)_j \leq M(1 - v_j), \quad \forall j \quad (2.135)$$

$$u \in \Omega_u, \quad \pi \geq 0, \quad w \in \{0, 1\}, \quad v \in \{0, 1\} \quad (2.136)$$

## 2.4 Summary

Mathematical theory based on interval computation and optimization is mainly introduced in this chapter. Interval computation includes interval linear equations, interval nonlinear equations; optimization methods based interval include interval optimization and robust optimization, where interval robust optimization include adaptive interval robust optimization and two-stage interval robust optimization. This chapter does not involve specific application model in power system, but use abstract mathematical expressions to describe and obtain some general solution methods. These methods provide the mathematical basis to solve the power system model created later in the paper, and can be widely used in other fields referring to interval models, such as water conservancy, transportation, aviation and so on. In addition, it notes that each mathematical model has its own solution which may be more than one, and the computational complexity or result precision may vary with each method. This paper contains many kinds of mathematical solution method which will be compared and discussed in the following application.

## References

1. Guessous N, Souhar O (2004) Recursive two-level ILU preconditioner for nonsymmetric M-matrices. *J Appl Math Comput* 16(1–2):19–35
2. Cui X, Hayami K (2009) Generalized approximate inverse preconditioners for least squares problems. *Jpn J Ind Appl Math* 26(1):1–14
3. Griffel DH (2002) *Applied functional analysis*. G. P. C. Press, Canada
4. Al-Khayyal FA, Larsen C, Van Voorhis T (1995) A relaxation method for nonconvex quadratically constrained quadratic programs. *J Global Optim* 6(3):215–230
5. Androulakis IP, Maranas CD, Floudas CA (1995)  $\alpha$ BB: a global optimization method for general constrained nonconvex problems. *J Global Optim* 7(4):337–363
6. Maranas CD, Floudas CA (1995) Finding all solutions of nonlinearly constrained systems of equations. *J Global Optim* 7(2):143–182
7. Raber U (1998) A simplicial branch-and-bound method for solving nonconvex all-quadratic programs. *J Global Optim* 13(4):417–432
8. Linderoth J (2005) A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. *Math Program* 103(2):251–282
9. Floudas CA, Pardalos PM, Nahanpetyan AG (2009) Bilinear programming. In: Floudas CA, Pardalos PM (eds) *Encyclopedia of optimization*. Springer, US, pp 279–282
10. Shapiro JF (1984) A note on node aggregation and Benders' decomposition. *Math Program* 29(1):113–119
11. Henry J, Yvon J, Lahdelma R et al (1994) A modified benders' decomposition technique for solving large scale unit commitment problems. In: Henry J, Yvon J (eds) *System modelling and optimization*. Springer, Berlin, pp 737–745
12. Floudas CA, Pardalos PM, Floudas CA (2001) Generalized benders decomposition. In: Floudas CA, Pardalos PM (eds) *Encyclopedia of optimization*. Springer, US, pp 747–758
13. Gass SI, Fu MC (2013) Benders decomposition method. In: Gass SI, Fu MC (eds) *Encyclopedia of operations research and management science*. Springer, US, p 115
14. Zeng B, Zhao L (2013) Solving two-stage robust optimization problems using a column-and-constraint generation method. *Oper Res Lett* 41(5):457–461

Power System Operation with Large Scale Stochastic  
Wind Power Integration  
Interval Arithmetic Based Analysis and Optimization  
Methods

Ding, T.

2017, XVII, 217 p. 70 illus., 64 illus. in color., Hardcover

ISBN: 978-981-10-2560-0