

## 2.2 Decomposition Optimization Framework

In this section we define the central notion of *safe decompositions*, and formulate our problem as one of finding the best feasible safe decomposition of a global constraint. The problem formulation in this section is applicable to all types of constraints.

### 2.2.1 Safe Decompositions

**Definition 1.** A constraint  $C$  is a Boolean function from the set of variables  $\vec{x}$ , to the Boolean set, i.e.,  $C: \text{Domain}(\vec{x}) \rightarrow \{\text{True}, \text{False}\}$

We denote the dimension (number of elements) in  $\vec{x}$  by  $|\vec{x}| = n$ .

**Definition 2.** A variable partition  $\mathbb{P}$  of the set of variables  $\vec{x}$  is defined as  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$ , such that  $\vec{y}_1 \cup \vec{y}_2 \cup \dots \cup \vec{y}_M = \vec{x}$ , and  $\vec{y}_i \cap \vec{y}_j = \emptyset$  for all  $i, j$  ( $1 \leq i, j \leq M, i \neq j$ ).

**Definition 3.** Let  $\Omega$  be a constraint, and  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$  be a partition of variables. We say that  $\mathbb{C} = (C_1, \dots, C_M)$  is a decomposition of  $\Omega$ , if in every constraint  $C_i$  all free variables are from  $\vec{y}_i$ . Sometimes we will use  $\mathbb{C}$  to indicate the conjunction  $C_1 \wedge \dots \wedge C_M$ . We say that a decomposition  $\mathbb{C} = (C_1, \dots, C_M)$  is safe if  $C_1 \wedge \dots \wedge C_M \models \Omega$ , where  $\models$  denotes logical entailment.

We also say that  $\mathbb{G} = (G_1, \dots, G_M)$  is a *cover* decomposition of  $\Omega$  if  $(G_1, \dots, G_M)$  is a decomposition of  $\Omega$  and  $\Omega \models G_1 \vee \dots \vee G_M$ <sup>7</sup>. The following proposition

---

<sup>7</sup>we will use  $\mathbb{G}$  to indicate the disjunction  $G_1 \vee \dots \vee G_M$ .

provides the relationship between safe and cover decompositions.

**Proposition 1.** *Let  $\Omega$  be constraint, and  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$  be a variable partition. Then,  $(C_1, \dots, C_M)$  is a safe (cover) decomposition of  $\Omega$  if and only if  $(\neg C_1, \dots, \neg C_M)$  is a cover (safe) decomposition of  $\neg\Omega$ .*

*Proof.* Since  $(C_1, \dots, C_M)$  is a safe (cover) decomposition of  $\Omega$ ,

$$\begin{aligned} C_1 \wedge \dots \wedge C_M \models \Omega &\Leftrightarrow \neg(\neg C_1 \vee \dots \vee \neg C_M) \models \Omega \\ &\Leftrightarrow \neg\Omega \models \neg C_1 \vee \dots \vee \neg C_M \end{aligned}$$

This completes the proof. □

In the following we will only concentrate on safe decompositions, but the results can also be applied to cover decompositions using Proposition 1.

**Definition 4.** *Let  $\vec{x}^0 = (\vec{y}_1^0, \dots, \vec{y}_M^0)$  be a database instance. We say that  $\vec{x}^0$  satisfies a safe decomposition  $\mathbb{C}$  if  $\vec{y}_1^0$  satisfies  $C_1$ ,  $\vec{y}_2^0$  satisfies  $C_2$ ,  $\dots$ , and  $\vec{y}_M^0$  satisfies  $C_M$ .*

**Example 1.** *Consider the following set of linear constraints:  $X + Y \leq 6$ ,  $-X + 5Y \leq 15$ ,  $5X + 4Y \leq 15$ , and both variables  $X$  and  $Y$  are non-negative. The partition  $\mathbb{P}$  is  $(\{X\}, \{Y\})$ , and a graphic representation of  $\Omega$  is given in Figure 2.1.*

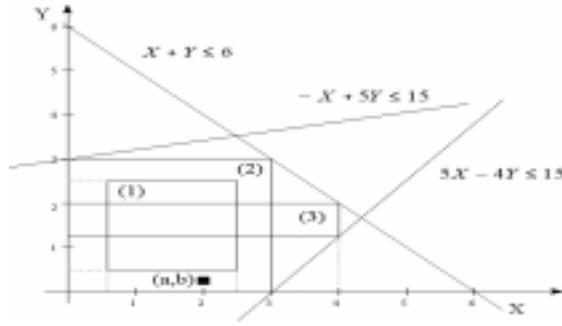


Figure 2.1: Safe Decompositions of  $\Omega$

Consider three safe decompositions  $\mathbb{C}_1, \mathbb{C}_2$ , and  $\mathbb{C}_3$ , where:

$$\begin{aligned} \mathbb{C}_1 &= (C_{11}, C_{12}) \\ &= (\{0.5 \leq X \leq 2.5\}, \{0.5 \leq Y \leq 2.5\}) \\ \mathbb{C}_2 &= (C_{21}, C_{22}), \text{ and} \\ &= (\{0.0 \leq X \leq 3.0\}, \{0.0 \leq Y \leq 3.0\}), \text{ and} \\ \mathbb{C}_3 &= (C_{31}, C_{32}) \\ &= (\{0.0 \leq X \leq 4.0\}, \{1.25 \leq Y \leq 2.0\}). \end{aligned}$$

$\mathbb{C}_1$  is a safe decomposition of  $\Omega$  because every point  $(X, Y)$  that satisfies  $C_{11}$  and  $C_{12}$  will also satisfy  $\Omega$ . Geometrically, this means that the space (1) defined by  $\mathbb{C}_1$  is contained in the space defined by  $\Omega$ . Similarly,  $\mathbb{C}_2$  and  $\mathbb{C}_3$  are safe decompositions of  $\Omega$ . Note that the database instance  $(a, b)$  satisfies  $\mathbb{C}_2$ , but not  $\mathbb{C}_1$  and  $\mathbb{C}_3$ .

Note that rectangle (1) (for  $\mathbb{C}_1$ ) is strictly contained in rectangle (2) (for  $\mathbb{C}_2$ ). Hence, the decomposition  $\mathbb{C}_2$  is better than  $\mathbb{C}_1$  in the sense that, in  $\mathbb{C}_2$  we will have to perform

global updates less frequently than in  $\mathbb{C}_1$ , i.e., less overhead. This notion is defined formally as follows:

**Definition 5.** *Given an arbitrary constraint  $\Omega$  and two decompositions  $\mathbb{C}_1 = (C_{11}, \dots, C_{1M})$  and  $\mathbb{C}_2 = (C_{21}, \dots, C_{2M})$ , we say that  $\mathbb{C}_2$  subsumes  $\mathbb{C}_1$  (or  $\mathbb{C}_1$  is subsumed by  $\mathbb{C}_2$ ) if:*

$$\bigwedge_{i=1}^M C_{1i} \models \bigwedge_{i=1}^M C_{2i}$$

*We will denote this by  $\mathbb{C}_1 \models \mathbb{C}_2$ . We say that  $\mathbb{C}_2$  strictly subsumes  $\mathbb{C}_1$  if  $\mathbb{C}_1 \models \mathbb{C}_2$ , but  $\mathbb{C}_2 \not\models \mathbb{C}_1$ . Furthermore, we say that a safe decomposition  $\mathbb{C}$  is minimally-constrained, if there is no safe decomposition  $\mathbb{C}'$  that strictly subsumes  $\mathbb{C}$ . Finally, we say that  $\mathbb{C}_1$  is equivalent to  $\mathbb{C}_2$ , denoted by  $\mathbb{C}_1 \equiv \mathbb{C}_2$ , if  $\mathbb{C}_1 \models \mathbb{C}_2$  and  $\mathbb{C}_2 \models \mathbb{C}_1$ .*

Note that, in example 1,  $\mathbb{C}_2$  and  $\mathbb{C}_3$  are minimally-constrained safe decompositions, while  $\mathbb{C}_1$  is not.

**Proposition 2.** *Let  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$  be a partition of variables, and  $\mathbb{C}$  be a conjunction of constraints  $(C_1, \dots, C_M)$ , where  $C_i$  ( $1 \leq i \leq M$ ) is over  $\vec{y}_i$ . Then,  $\mathbb{C}$  is satisfiable iff for every  $i$ ,  $1 \leq i \leq M$ ,  $C_i$  is satisfiable.*

*Proof.* IF-part, if every  $C_i$  is satisfiable, its conjunction is satisfiable, i.e.,  $\mathbb{C}$  is satisfiable. ONLY-IF-part, if  $\mathbb{C}$  is satisfiable, and since (1)  $(\vec{y}_1, \dots, \vec{y}_M)$  is a partition of  $\vec{x}$ , and (2) each  $C_i$ ,  $1 \leq i \leq M$ , has its free variables exclusively in  $\vec{y}_i$ , then, all  $C_i$ 's are satisfiable. □

**Proposition 3.** *Let  $\mathbb{C}_1 = (C_{11}, \dots, C_{1M})$  and  $\mathbb{C}_2 = (C_{21}, \dots, C_{2M})$  be two lists of constraints over  $\vec{y}_1, \dots, \vec{y}_M$  respectively, for partition  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$ . Then:*

1. *If  $\mathbb{C}_1$  is satisfiable, then  $\mathbb{C}_1 \models \mathbb{C}_2$  iff for all  $i$ ,  $1 \leq i \leq M$ ,  $C_{1i} \models C_{2i}$ .*
2. *If both  $\mathbb{C}_1$  and  $\mathbb{C}_2$  are satisfiable  $\mathbb{C}_1 \equiv \mathbb{C}_2$  iff for all  $i$ ,  $1 \leq i \leq M$ ,  $C_{1i} \equiv C_{2i}$ .*

*Proof.* Part 2 immediately follows from Part 1. In Part 1, the " $\Leftarrow$ " direction is obvious, while the " $\Rightarrow$ " is due the fact that the variable partition  $\mathbb{P} = (\vec{y}_1, \dots, \vec{y}_M)$  is disjoint as follows:

Assume that  $\mathbb{C}_1 \models \mathbb{C}_2$ , but, by way of contradiction, for some  $i$ ,  $1 \leq i \leq M$ ,  $C_{1i} \not\models C_{2i}$ . Then, there exists  $\vec{a}_i$  over  $\vec{y}_i$  that satisfies  $C_{1i}$ , but not  $C_{2i}$ . Since,  $\mathbb{C}_1$  is consistent, each  $C_{11}, \dots, C_{1M}$  must be consistent, and therefore there must exist  $\vec{b}_1, \dots, \vec{b}_M$  over  $\vec{y}_1, \dots, \vec{y}_M$ , that satisfy  $C_{11}, \dots, C_{1M}$ , respectively. Then,  $\vec{b}_1, \dots, \vec{b}_{i-1}, \vec{a}_i, \vec{b}_{i+1}, \dots, \vec{b}_M$  satisfies  $\mathbb{C}_1$ , but not  $\mathbb{C}_2$ , contradicting the fact that  $\mathbb{C}_1 \models \mathbb{C}_2$ .  $\square$

In practical cases, we are only interested in the case when  $\Omega$  is satisfiable, because otherwise the database must be empty and no update would be allowed. Technically, however, every unsatisfiable (i.e., inconsistent) decomposition will be safe for unsatisfiable  $\Omega$ . If  $\Omega$  is satisfiable we have the following:

**Proposition 4.** *Let  $\Omega$  be a satisfiable constraint. Then every minimally-constrained safe decomposition  $\mathbb{C}$  of  $\Omega$  is satisfiable.*

*Proof.* Since  $\Omega$  is satisfiable, there exists  $\vec{a} = (\vec{a}_1, \dots, \vec{a}_M)$  over  $\vec{x}$  that satisfies  $\Omega$ .

Then, the decomposition  $\mathbb{C}_1 = (\vec{y}_1 = \vec{a}_1, \dots, \vec{y}_M = \vec{a}_M)$  is always a safe decomposition of  $\Omega$ .

Consider now an arbitrary minimally-constrained safe decomposition  $\mathbb{C}$ . If, by way of contradiction,  $\mathbb{C}$  is not satisfiable, then  $\mathbb{C} \models \mathbb{C}_1$  and  $\mathbb{C}_1 \not\models \mathbb{C}$ , contradicting the minimality of  $\mathbb{C}$ .  $\square$

Clearly, safe or even minimally-constrained safe decompositions are not unique. In our example, both  $\mathbb{C}_2$  and  $\mathbb{C}_3$  are minimally-constrained, because there is no other safe decomposition that strictly subsumes  $\mathbb{C}_2$  or  $\mathbb{C}_3$ .

Since safe decompositions are not unique, an important question is how to choose a safe decomposition that is optimal according to some meaningful criterion.

In our example, the rectangle with the maximum area may be a good choice. In fact, if update points  $(X,Y)$  are uniformly distributed over the given space (defined by  $\Omega$ ), then the larger area (volume in the general case) corresponds to greater probability that an update will satisfy local constraints, and thus no global processing will be necessary. We defer the discussion on optimality criteria to Section 2.6.

## 2.2.2 Optimization Problem Formulation

We suggest the following general framework for selecting optimal feasible decompositions:

$$\begin{aligned}
& \text{maximize } f(s) \\
& \text{s.t. } s \in \mathbb{S}
\end{aligned} \tag{2.1}$$

where  $\mathbb{S}$  is the set of all feasible decompositions, and  $f : \mathbb{S} \rightarrow \mathbb{R}$  (real numbers) is the objective function discussed in the next subsection.

**Definition 6.** *Let  $\Omega$  be a global constraint,  $\mathbb{C} = (C_1, \dots, C_M)$  be a decomposition of  $\Omega$ ,  $\theta = \{k+1, \dots, M\}$  be a subset of sites  $\{1, \dots, M\}$ , and  $\vec{x} = (\vec{y}_1, \dots, \vec{y}_M)$  be a database instance. Then, feasible decompositions are defined as the set of decompositions having one or more of the following properties:*

1. *Safety. Decomposition  $\mathbb{C}$  has the safety property if  $\mathbb{C}$  is a safe decomposition of  $\Omega$ .*
2. *Local Consistency. Decomposition  $\mathbb{C}$  has the local consistency property if each local instance  $\vec{y}_i$  satisfies its local constraint  $C_i$  ( $1 \leq i \leq M$ ). Clearly, local consistency and safety imply global consistency.*
3.  *$\theta$ -Partial Constraint Preservation. Decomposition  $\mathbb{C}$  has the partial constraint preservation w.r.t.  $\theta$ , if local constraints for sites outside of  $\theta$  are fixed. This property can be used for scenarios where a (safe) constraint re-decomposition would involve only a (hopefully small) subset of sites  $\theta$ , leaving the rest unchanged.*

4.  $\theta$ -Resource Bound Partition <sup>8</sup>. *Decomposition*  $\mathbb{C}$  has the resource bound partition property w.r.t.  $\theta$  and a bound  $\vec{B}_\theta$ , if the overall resource of sites in  $\theta$  is bounded by  $\vec{B}_\theta$ .

Resource bound partition property is for families of constraints in which resource specification is possible (such as linear constraints). Namely, the global constraint  $\Omega$  is associated with the global resource bound  $\vec{b}$ , each local constraint  $C_i$  in the decomposition is associated with a resource  $\vec{r}_i$ , and each subset of sites  $\theta$  is associated with the cumulative resource  $\vec{r}_\theta$ .

A partition of the global resource bound between  $\theta$  and  $\bar{\theta}$  (i.e., all sites except  $\theta$ ) is a pair  $(\vec{B}_\theta, \vec{B}_{\bar{\theta}})$ , such that  $\vec{B}_\theta + \vec{B}_{\bar{\theta}} = \vec{b}$ , is a partition identified by  $\vec{B}_\theta$ . We say that a decomposition  $\mathbb{C}$  has a *resource bound partition* property w.r.t. a partition  $\vec{B}_\theta$ , if the cumulative resource  $\vec{r}_\theta$  is bounded by  $\vec{B}_\theta$ , and  $\vec{r}_{\bar{\theta}}$  is bounded by  $\vec{B}_{\bar{\theta}}$ .

Before we discuss how the decomposition problems can be solved effectively, we consider possible candidates for function  $f$ .

### 2.2.3 Optimization Criteria

There are many feasible (minimally-constrained) safe decompositions in  $\mathbb{S}$ , and we would like to formulate a criterion to select the best among them. This criterion should represent the problem characteristics, and the decomposition goals. Possible

---

<sup>8</sup>Note that the notion of resource bound partition is more flexible than constraint preservation, and allows one to perform concurrent constraint decompositions.

criteria include:

- Maximize the probability that an update will not violate the existing local constraints (decomposition).
- Minimize overall expected cost of computations during an update.
- Maximize the expected number of updates before the first update that violates local constraints.
- Maximize the expected length of time before an update violates local constraints.

Many other optimization criteria are possible. However, any reasonable criteria should be monotonic, as defined below.

**Definition 7.** *Let  $f$  be a function from the set of safe decompositions of  $\Omega$  to  $\mathbb{R}$ . We say that  $f$  is monotonic if for every two decompositions  $\mathbb{C}_1, \mathbb{C}_2$  of  $\Omega$ ,  $\mathbb{C}_1 \models \mathbb{C}_2$  implies  $f(\mathbb{C}_1) \leq f(\mathbb{C}_2)$ .*

Intuitively, being monotonic for an optimization criterion means that enlarging the space defined by a decomposition can only make it better.

Note, that if  $f$  is monotonic, then  $f(\mathbb{C}_1) = f(\mathbb{C}_2)$  for any two equivalent decompositions  $\mathbb{C}_1$  and  $\mathbb{C}_2$ .

As we will see in Section 2.5, it is often necessary to consider a subspace of all safe decompositions (without losing an optimal decomposition).

**Definition 8.** Let  $\mathbb{S}$  be a set of safe decompositions of  $\Omega$ . A subset  $\mathbb{S}'$  of  $\mathbb{S}$  will be called a *monotonic cover* of  $\mathbb{S}$  if for every decomposition  $\mathbb{C}$  in  $\mathbb{S}$  there exists a decomposition  $\mathbb{C}'$  in  $\mathbb{S}'$ , such that  $\mathbb{C}'$  subsumes  $\mathbb{C}$  (i.e.,  $\mathbb{C} \models \mathbb{C}'$ ).

The following proposition states that optimal decompositions are not missed when the search space is restricted to a monotonic cover and the optimization criteria are monotonic.

**Proposition 5.** Let  $\mathbb{S}$  be a (sub) set of all safe decompositions of  $\Omega$ ,  $\mathbb{S}'$  be a monotonic cover of  $\mathbb{S}$ , and  $f$  be a monotonic function from  $\mathbb{S}$  to  $\mathbb{R}$ . Then, the following two optimization problems yield the same maximum.

**Problem 1.**  $\max f(s), \text{ s.t. } s \in \mathbb{S}$ .

**Problem 2.**  $\max f(s), \text{ s.t. } s \in \mathbb{S}'$ .

*Proof.* Suppose the maxima of  $f$  in Problem 1 and 2 are achieved by  $s = \mathbb{C}$  in  $\mathbb{S}$  and  $s = \mathbb{C}'$  in  $\mathbb{S}'$  respectively. Since  $\mathbb{S}' \subseteq \mathbb{S}$ ,  $f(\mathbb{C}') \leq f(\mathbb{C})$ . Now, since  $\mathbb{S}'$  is a monotonic cover of  $\mathbb{S}$ , there must exist  $\mathbb{C}'' \in \mathbb{S}'$ , such that  $\mathbb{C} \models \mathbb{C}''$ . Therefore, because  $f$  is monotonic,  $f(\mathbb{C}) \leq f(\mathbb{C}'')$ .

Finally, since the maximum of Problem 2 is achieved at  $\mathbb{C}'$ ,  $f(\mathbb{C}') \geq f(\mathbb{C}'') \geq f(\mathbb{C})$ . Thus,  $f(\mathbb{C}') = f(\mathbb{C})$  which completes the proof.  $\square$

The functions in the above criteria depend on the update distribution and other assumptions. Specifically we consider two assumptions.

- For design-time decompositions: We do not know the current database state, but we are given a probability distribution of database instances in the space defined by  $\Omega$ .
- For update-time decompositions: We are given a current database instance, and a conditional distribution function of database instances on  $\Omega$ .

Now, we present precise methods to characterize the set  $\mathbb{S}$ , function  $f$ , and algorithms to solve effectively the optimization problem (2.1), for the family of linear constraints.

## 2.3 Linear Arithmetic Constraints

**Definition 9.** *An atomic linear constraint is an inequality of the form  $a_1x_1 + a_2x_2 + \dots + a_nx_n \leq b$ , where  $a_1, a_2, \dots, a_n$ , and  $b$  are real numbers, and  $x_1, x_2, \dots, x_n$  are variables ranging over the reals.*

Definition 9 defines a constraint as a symbolic expression. However, an atomic linear constraint  $a_1x_1 + a_2x_2 + \dots + a_nx_n \leq b$  also defines a Boolean function  $C : \mathbb{R}^n \rightarrow \{True, False\}$ , where for each instantiation of values to variables  $(x_1, x_2, \dots, x_n)$ , the expression  $a_1x_1 + a_2x_2 + \dots + a_nx_n \leq b$  is evaluated as true or false.

**Definition 10.** *A linear system  $\Omega$  is a conjunction of atomic linear constraints.*

A linear system  $\Omega$  containing  $n$  variables and (a conjunction of)  $m$  atomic linear constraints, can be written as follows:

$$\begin{array}{cccccc}
 a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & \leq & b_1 \\
 a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{2n}x_n & \leq & b_2 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \\
 a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & \leq & b_m
 \end{array} \tag{2.2}$$

This system  $\Omega$  can also be written in matrix notation as the system  $A\vec{x} \leq \vec{b}$ , where  $A$  is the matrix

$$\begin{pmatrix}
 a_{11} & a_{12} & \dots & a_{1n} \\
 a_{21} & a_{22} & \dots & a_{2n} \\
 \vdots & \vdots & \vdots & \vdots \\
 a_{m1} & a_{m2} & \dots & a_{mn}
 \end{pmatrix} \tag{2.3}$$

and  $\vec{b}$  is the column vector  $(b_1 \ b_2 \ \dots \ b_m)$ , and  $\vec{x}$  is the vector  $(x_1 \ x_2 \ \dots \ x_n)$ .

### 2.3.1 Parametric Optimization Problem

To solve the optimization problem for linear arithmetic constraints we want to rewrite the problem (2.1), i.e.,

$$\begin{aligned} \max f(s) \\ \text{s.t. } s \in \mathbb{S} \end{aligned}$$

where  $\mathbb{S}$  is the set of feasible safe decompositions in the form

$$\begin{aligned} \max f(\vec{w}) \\ \text{s.t. } \Phi(\vec{w}) \end{aligned}$$

where  $\vec{w}$  is the set of variables describing coefficients (i.e., parameters) of constraints on a decomposition  $D(\vec{w})$ , and  $\Phi(\vec{w})$  is a logical condition in terms of  $\vec{w}$  defining the search space

$$\mathbb{S}' = \{D(\vec{w}) \mid \Phi(\vec{w})\}$$

such that  $\mathbb{S}'$  is a monotonic cover of  $\mathbb{S}$ .

By Proposition 5, the two problems are equivalent for any user-given monotonic optimization function, but the latter allows the use of known mathematical programming methods to solve it.

We do it for the case of an individual as well as a general variable partition as

described in Section 2.4 and 2.5, respectively, in which we study the problem of parametric characterization of decompositions.

## 2.4 Individual Variable Partitions

Individual partition case variables are partitioned in an individual way, i.e., we have a partition  $\mathbb{P} = (\{x_1\}, \dots, \{x_n\})$ , where  $\{x_1, x_2, \dots, x_n\}$  is the set of all variables<sup>9</sup>.

### 2.4.1 Parametric Characterization

In this case, safe decompositions can be parametrically described using intervals as follows:

**Proposition 6.** *Given a bounded set of constraints  $\Omega$ , and an individual partition  $\mathbb{P}$ , every decomposition  $\mathbb{C} = (C_1, \dots, C_n)$  of  $\Omega$  is equivalent to a decomposition  $\mathbb{C}'$  of  $\Omega$  of the form:*

$$(\{u_{11} \leq x_1 \leq u_{21}\}, \dots, \{u_{1n} \leq x_n \leq u_{2n}\}) \quad (2.4)$$

*Proof.* Every atomic constraint  $C_i$  ( $1 \leq i \leq n$ ) of  $\mathbb{C}$  over  $\vec{x}$  can be written as  $x_i \leq v_{ij}$  or  $z_{ij} \leq x_i$ . Thus  $C_i$  will be equivalent to  $u_{1i} \leq x_i \leq u_{2i}$ , where  $u_{1i} = \max\{z_{i1}, z_{i2}, \dots, z_{in_i}\}$  and  $u_{2i} = \min\{v_{i1}, v_{i2}, \dots, v_{in_i}\}$ . □

---

<sup>9</sup>Note that this implies  $n = M$ .

In this section, we will denote the decomposition  $(\{u_{11} \leq x_1 \leq u_{21}\}, \dots, \{u_{1n} \leq x_n \leq u_{2n}\})$  by  $D(\vec{u})$ , where  $\vec{u} = (\vec{u}_2, \vec{u}_1)$ ,  $\vec{u}_1 = (u_{11}, u_{12}, \dots, u_{1n})$  and  $\vec{u}_2 = (u_{21}, u_{22}, \dots, u_{2n})$ . We use the notation  $\vec{u}_1 \leq \vec{u}_2$  to denote that  $u_{1i} \leq u_{2i}$ , for all  $i$ ,  $1 \leq i \leq n$ .

To create a parametric characterization of the set of all safe decompositions we introduce the notion of characterization matrix as follows.

**Definition 11.** *Given an  $n \times m$  matrix  $A$ , the characterization matrix  $A'$  of  $A$  is defined as  $(A^+ A^-)$ , where both are  $n \times m$  matrices with elements  $a_{ij}^+$  and  $a_{ij}^-$  respectively, defined as follows:*

$$a_{ij}^+ = \begin{cases} a_{ij} & \text{if } a_{ij} > 0, \\ 0 & \text{otherwise} \end{cases}$$

$$a_{ij}^- = \begin{cases} a_{ij} & \text{if } a_{ij} < 0, \\ 0 & \text{otherwise} \end{cases}$$

In fact, we have developed the notion of characterization matrix so that the following holds:

**Theorem 1.** *Let  $S$  be the system  $A\vec{x} \leq \vec{b}$ . Let  $A'$  be the characterization matrix of  $A$ , and  $S'$  be the system  $A'\vec{u} \leq \vec{b}$  and  $\vec{u}_1 \leq \vec{u}_2$ , where  $\vec{u} = (\vec{u}_2, \vec{u}_1)$ . Then, for every  $\vec{u}_1^0, \vec{u}_2^0$  the following are equivalent:*

1.  $\vec{u}^0$  is a solution of  $S'$ , i.e.,  $A'\vec{u}^0 \leq \vec{b}$  and  $\vec{u}_1^0 \leq \vec{u}_2^0$

2. Every  $\vec{v}$ ,  $\vec{u}_1^0 \leq \vec{v} \leq \vec{u}_2^0$ , is a solution of  $S$ , i.e.,  $A\vec{v} \leq \vec{b}$

*Proof.* First, we prove (1)  $\Rightarrow$  (2). Assume that  $\vec{u}^0 = (\vec{u}_1^0, \vec{u}_2^0)$  is a solution of  $S'$ . Let  $\vec{v}$  be a vector such that  $u_{1i}^0 \leq v_i \leq u_{2i}^0$  for all  $i$ ,  $1 \leq i \leq n$ . By multiplying  $v_i \leq u_{2i}^0$  by a non-negative number  $\alpha_i^+$ , we get  $\alpha_i^+ v_i \leq \alpha_i^+ u_{2i}^0$  for all  $i$ ,  $1 \leq i \leq n$ . Now, choosing  $\alpha_i^+$ 's as the elements of  $j^{\text{th}}$  column of  $A^+$ , and making a summation for all possible elements in that row, we get:

$$\sum_{i=1}^n a_{ji}^+ v_i \leq \sum_{i=1}^n a_{ji}^+ u_{2i}^0 \quad (2.5)$$

and extending for all possible rows in  $A^+$ , we get:

$$A^+ \vec{v} \leq A^+ \vec{u}_2^0 \quad (2.6)$$

Repeating the same operations for  $u_{1i}^0 \leq v_i$ , i.e., multiplying by a non-positive number  $\alpha_i^-$ , we get:  $\alpha_i^- u_{1i}^0 \geq \alpha_i^- v_i$  for all  $i$ ,  $1 \leq i \leq n$ . Now, choosing  $\alpha_i^-$ 's as the elements of  $j^{\text{th}}$  column of  $A^-$ , and making a summation for all possible elements in that row, we get:

$$\sum_{i=1}^n a_{ji}^- u_{1i}^0 \geq \sum_{i=1}^n a_{ji}^- v_i \quad (2.7)$$

and extending for all possible rows in  $A^-$ , we get:

$$A^- \vec{u}_1^0 \geq A^- \vec{v} \quad (2.8)$$

Now, adding (2.6) and (2.8) we get:

$$A^+ \vec{v} + A^- \vec{v} \leq A^+ \vec{u}_2^0 + A^- \vec{u}_1^0 \quad (2.9)$$

We know that  $\vec{u}^0 = (\vec{u}_2^0, \vec{u}_1^0)$  is a solution of  $A' \vec{u} \leq \vec{b}$ , then:

$$A^+ \vec{v} + A^- \vec{v} \leq \vec{b} \text{ or} \quad (2.10)$$

$$(A^+ + A^-) \vec{v} \leq \vec{b}$$

Therefore, by definition of  $A^+$  and  $A^-$ ,  $A \vec{v} \leq \vec{b}$ , i.e.,  $\vec{v}$  is a solution of  $S$ , which completes the proof (1)  $\Rightarrow$  (2).

We prove now that (2)  $\Rightarrow$  (1), by proving  $\neg$  (1)  $\Rightarrow$   $\neg$  (2). Assume that  $\vec{u}^0$  is not a solution of  $S'$ , i.e., vector  $\vec{u}^0$  does not satisfy  $A' \vec{u} \leq \vec{b}$ . Then, there exists a column  $j$  such that:

$$\sum_{i=1}^n a_{ij}^+ u_{2i}^0 + \sum_{i=1}^n a_{ij}^- u_{1i}^0 > b_j \quad (2.11)$$

Consider a vector  $\vec{v}$  defined as follows:

$$v_i = \begin{cases} u_{2i}^0 & \text{if } a_{ij}^+ > 0, \\ u_{1i}^0 & \text{if } a_{ij}^- < 0, \\ u_{2i}^0 & \text{if } a_{ij}^+ = a_{ij}^- = 0 \end{cases}$$

Clearly,  $\vec{u}_1^0 \leq \vec{v} \leq \vec{u}_2^0$ . Then, we can rewrite (2.11) as

$$\sum_{i=1}^n a_{ij} v_i > b_j \quad (2.12)$$

Therefore,  $\vec{v}$  is not a solution to  $A\vec{x} \leq \vec{b}$ . This completes the proof.  $\square$

## 2.4.2 Parametric Optimization Problem

We are now ready to characterize parametrically the optimization problem of safe decompositions of  $\Omega$ .

**Proposition 7 (Parametric Feasible Properties).** *Let  $\Omega = A\vec{x} \leq \vec{b}$  be a global satisfiable constraint,  $\mathbb{P} = (x_1, \dots, x_n)$  be an individual variable partition of  $\vec{x}$ ,  $A'$  the characterization matrix of  $A$ ,  $\vec{x}^0 = (x_1^0, \dots, x_n^0)$  be an instance of  $\vec{x}$ ,  $\theta$  be a subset  $\{k + 1, \dots, M\}$  of sites  $\{1, \dots, M\}$  and  $\bar{\theta}$  be its complement, and  $D(u_{11}^0, u_{21}^0, \dots, u_{1k}^0, u_{2k}^0)$  be a (partial)  $\bar{\theta}$ -safe decomposition that satisfies  $(x_1^0, \dots, x_k^0)$ . Then, for any decomposition  $D(u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n})$*

1.  $D(u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n})$  is safe iff  $A'\vec{u} \leq \vec{b}$  and  $\vec{u}_1 \leq \vec{u}_2$ . We denote this condition by  $\Phi_{safe}(\vec{u})$ .
2.  $D(u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n})$  satisfies local consistency w.r.t.  $\vec{x}^0$  iff for all  $i$ ,  $1 \leq i \leq n$ ,  $u_{1i} \leq x_i^0 \leq u_{2i}$ . We denote this condition by  $\Phi_{lc}(\vec{u})$ .
3.  $D(u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n})$  satisfies partial constraint preservation w. r.t.  $\bar{\theta}$ -safe decomposition  $D(u_{11}^0, u_{21}^0, \dots, u_{1k}^0, u_{2k}^0)$  iff  $u_{11} = u_{11}^0, u_{21} = u_{21}^0, \dots, u_{1k} = u_{1k}^0, u_{2k} = u_{2k}^0$ . We denote this condition by  $\Phi_{pcp}(\vec{u})$ .

*Proof.* 1) follows directly from Theorem 1 and Proposition 6. IF-part: if  $A'\vec{u} \leq \vec{b}$  and  $\vec{u}_1 \leq \vec{u}_2$ , by Theorem 1 (i.e., (1)  $\Rightarrow$  (2)),  $D(\vec{u})$  is safe. ONLY-IF-part: if  $D(u_{11}, \dots, u_{1n}, u_{21}, \dots, u_{2n})$  is safe, i.e.,  $\vec{u}_1 \leq \vec{u}_2$ . Since  $D(\vec{u})$  is a safe decomposition, the condition (2) of Theorem 1 holds, and thus the condition (1), namely  $A'\vec{u} \leq \vec{b}$  and  $\vec{u}_1 \leq \vec{u}_2$ , which completes this part of the proof.

Now, 2) follows from the definition of local consistency, and 3) This follows directly from partial constraint preservation definition.  $\square$

We denote by  $\mathbb{S}_{safe}$ ,  $\mathbb{S}_{lc}$ ,  $\mathbb{S}_{pcp}$  the set of safe decompositions, the set of decompositions satisfying local consistency w.r.t.  $\vec{x}^0$ , and decompositions satisfying partial constraint preservation w.r.t. a  $\theta$ -safe decomposition  $D(u_{11}^0, u_{21}^0, \dots, u_{1k}^0, u_{2k}^0)$  respectively. We will use  $Pr$  to denote a subset of the set of properties  $\{safety, lc, pcp\}$ . Finally, set  $\mathbb{S}_{Pr}$  will denote the set of all decompositions that satisfy the properties  $Pr$ , i.e.,  $\mathbb{S}_{Pr} = \bigcap_{p \in Pr} \mathbb{S}_p$ , and  $\Phi_{Pr}(\vec{u})$  will be the conjunction of the corresponding conditions, i.e.,

$\Phi_{Pr}(\vec{u}) = \bigwedge_{p \in Pr} \Phi_p(\vec{u})$ . We can present the optimization problem in terms of resource characterization.

**Theorem 2 (Parametric Optimization Problem).** *Let  $\Omega = A\vec{x} \leq \vec{b}$  be a satisfiable global constraint,  $f$  be a monotonic function from the set of all safe decompositions to  $\mathbb{R}$ ,  $\mathbb{P} = (\vec{x}_1, \dots, \vec{x}_n)$  a individual variable partition of  $\vec{x}$ ,  $\vec{x}^0 = (\vec{x}_1^0, \dots, \vec{x}_n^0)$  be an instance of  $\vec{x}$ ,  $D(u_{11}^0, u_{21}^0, \dots, u_{1k}^0, u_{2k}^0)$ ,  $1 \leq k \leq M$ , a partial safe decomposition that satisfies  $(x_1^0, \dots, x_k^0)$ , and  $Pr$  be the subset of properties  $\{\text{safety}, \text{lc}, \text{pcp}\}$  that must contain safety. Then, solving the optimization problem*

$$\begin{aligned} \max f(s) \\ \text{s.t. } s \in \mathbb{S}_{Pr} \end{aligned}$$

*is equivalent to solving the parametric problem*

$$\begin{aligned} \max f(D(\vec{u})) \\ \text{s.t. } \Phi_{Pr}(\vec{u}) \end{aligned}$$

*Proof.* First, by Proposition 6, every decomposition  $\mathbb{C}$  in  $\mathbb{S}_{Pr}$  has an equivalent decomposition  $\mathbb{C}'$  of the form  $D(\vec{u}^0)$ , where  $\vec{u}_1^0 \leq \vec{u}_2^0$ . Then, by Proposition 7,  $\mathbb{S}_{Pr}$  and  $\Phi_{Pr}(\vec{u})$  represent equivalent search spaces. Therefore, both problems yield the same maximum. This completes the proof.  $\square$