AN OPTIMAL REPLACEMENT - DESIGN MODEL FOR A RELIABLE WATER DISTRIBUTION NETWORK SYSTEM

by

Ernest Price Smith

Dissertation submitted to the faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial and Systems Engineering

APPROVED:

H. D. Sherali, Chairman

G. V. Loganathan

R. F. Olin

J. A. Nachlas

S. C. Sarin

October 1994

Blacksburg, Virginia

An Optimal Replacement - Design Model
for a Reliable Water Distribution Network System

by

Ernest P. Smith
Dr. Hanif D. Sherali, Chairman
Industrial and Systems Engineering

(ABSTRACT)

A municipal water distribution system is a network of underground pipes, usually mirroring the city street network, that connects water supply sources such as reservoirs and water towers with demand points such as residential homes, industrial sites and fire hydrants. These systems are extremely expensive to install, with costs typically in the tens of millions of dollars. Existing pipes wear out, leak and break due to various factors including corrosion and cold weather ruptures, each requiring inconvenient and expensive repairs. Therefore, over time, these pipes need to be considered for replacement. Meanwhile, increasing urban development and water use rates dictate the need for larger and larger pipes to handle increased flow rates, while keeping water pressure within an acceptable range. However, larger pipes cost considerably more to install and maintain, causing the pipe sizing decision to be a critical task.

We develop an optimal network design and replacement strategy that meets hydraulic requirements under all likely demand and failure scenarios. Two submodels are used in a hierarchical fashion to integrate the reliability and cost analysis, and the network optimization process, within the overall network design process. The pipe reliability and cost submodel uses statistical methodologies based on historical records of pipe breaks to estimate future maintenance costs, and to recommend replacing relatively expensive-to-maintain or undercapacitated pipes. The pipe network optimization submodel provides a least cost construction and replacement plan along with optimal flows and energy heads for each fixed network configuration and demand pattern.

Traditional approaches isolate the above two types of models, assuming away the required interaction of inputs and outputs between them. We use a hierarchical design approach that integrates the foregoing two submodels by designing the network in a sequential fashion over a number of stages. The models are tied in a feedback loop that reprocesses the information until a stable design is attained. The result is a comprehen-

sive reduced cost network design that meets all pressure and flow requirements for realistic problems, even under a wide variety of pipe failure modes.

For the optimization model, we develop two new algorithms that exploit the special network structure of the problem. In the first approach, the problem is restructured in a manner that facilitates its decomposition into a master control program, and an easy-to-solve convex cost network flow programming subproblem. The master program operates in the space of the structural design variables, while the subproblem determines flows as well as heads via its primal and dual optimal solutions. The coordination between the master program and the subproblems is effected via a suitable penalty term. The theoretical validity of the decomposition scheme is established, and efficient algorithmic implementation strategies are developed. On a standard popular test problem in the literature, this procedure is shown to recover a solution that significantly improves upon a previously best known solution.

The second optimization approach is one that guarantees a global optimal solution, and contrasts with previous approaches that at best produce local optimal solutions. This procedure is based on a Reformulation-Linearization Technique that constructs tight linear programming relaxations for the nonlinear problem, and embeds these in a branch-and-bound algorithm. A suitable partitioning strategy is coordinated with this scheme to provably ensure infinite convergence to a global optimum. When this method is applied to the aforementioned test problem, a further improved solution is obtained.

It is hoped that with additional enhancements and refinements, our proposed methodologies will serve to provide a useful tool for practitioners to design reliable pipe network water distribution systems.

Acknowledgements

I would like to thank the members of my committee for their assistance in preparing this document and for tolerating my last minute preparations and schedule limitations. Special thanks go to Dr. Sherali for his patient and careful editing, feedback and insightful instruction. Many thanks also to my friends and family who helped me through this difficult time. This document is dedicated to my kids, Stephen and Rachel, who gave up so many hours with their dad so that this dissertation could be completed. I recommend that you get all your education finished before having kids of your own!

Table of Contents

1. Introduction and Motivation

1.1 Introduction 1

1.2 Practical Considerations 2

1.3 Proposed Approach 7

2. Literature Review

2.1 Pipe Reliability and Cost Models 9

2.2 Pipe Network Analysis Models 14

2.3 Single Demand Pipe Network Design Models 16

2.4 Integrated Pipe Network Design Models 23

2.5 Reformulation-Linearization Technique (RLT) 26

3. Model Formulations and Integrated Approach

3.1 Integrated Design Approach 28

3.2 Pipe Network Design Submodel Formulation 30

3.2.1 Network Optimization Problem Formulation 33

3.3 Pipe Reliability and Cost Submodel Formulation 38

3.4 Model Parameters 54

4. Decomposition Algorithm for the Network Optimization Problem

4.1 Structural Manipulation and Reformulation of the Network Optimization Problem 55

4.2 A Dimension Reduction Via the Definition of Composite Variables for PNOP 63

4.3 The Projected Network Optimization Decomposition Algorithm 67

4.4 Algorithm for Solving the Subproblem SP 71

4.5 Algorithm for Estimating Gradients Via a Barrier Subproblem BSP 73

4.6 Computing Gradients for PNOP 74

5. A Global Optimization Approach for the Network Design Problem

5.1 Introduction to the RLT Methodology 80
5.2 Reformulation Phase 84
5.3 Linearization Phase 86
5.4 Branch-and-Bound Strategy 93
5.5 Convergence of the RLT Branch-and-Bound Algorithm 96
5.6 Size of the Problem RLTNOP 99
5.7 Practical Considerations 99
5.8 Modifications for Rational Exponents 100

6. Computational Experience on a Standard Test Problem

6.1 Results from the Decomposition Algorithm of Chapter 4 103
6.2 Results from the RLT Global Optimization Algorithm of Chapter 5 106
6.3 RLT Results for the Design Problem Having Integer Exponents in the Hazen-Williams Equations 110

7. Summary, Conclusions and Recommendations for Future Research

7.1 Summary and Significance of Results 112
7.2 Recommendations for Future Research 114

Bibliography 117

VITA 123

1. INTRODUCTION AND MOTIVATION

1.1 Introduction

The quality of water distribution systems (WDS) plays a crucial role in a society because of its strong contribution to community health, firefighting capability, quality of life, and the potential for future economic growth. Aging, deteriorating water systems have long been recognized as a critical national problem. Many of these systems are old, and pipes are heavily tuberculated or undersized for the present day use, and hence unable to provide the required discharge at adequate pressure heads. Pipe failures are a daily occurrence in many cities, and recent catastrophic failures in Chicago and Washington, D.C., and impending fears with respect to the New York City distribution network have raised an awareness of the need for upgrading the water distribution infrastructure to an acceptable level of serviceability and reliability. Meanwhile, new sectors are being continually added to existing networks in order to accommodate growing communities and new industries. The costs associated with the overhaul and expansion of these systems are prohibitive and can easily escalate into the tens of millions of dollars.

As traditional methods for designing WDS are heuristic in nature and incorporate high levels of redundancy, they tend to be overly conservative, and therefore expensive. Savings of 10-30% have regularly been accomplished through the use of computerized algorithms, but most approaches have not adequately addressed the high redundancy requirements characterized by these systems. Therefore, one needs to accurately assess the existing condition of a given WDS network, and to prescribe a cost effective construction plan for its renovation and expansion, while retaining or designing for a high level of reliability and redundancy.

The most general problem is to modify and/or expand the design of an existing network by prescribing high enough energy heads and large enough pipe diameters of sufficient smoothness to supply the varied anticipated demand for water at required pressure levels, even while experiencing breakages in the network. If the network is designed with low energy heads and undersized or rough pipes in a skeletal fashion, then flow and/or pressure requirements will not be met during certain demand peaks or under various pipe failure scenarios. On the other hand, if the energy sources and pipes are overdesigned, or if there are too many redundant paths, then increased costs may lead to an inefficient solution. Therefore, there is a strong need for determining an optimal network design and replacement strategy that meets hydraulic requirements under all demand patterns and likely failure modes. This is the principal consideration in the present study.

For clarification, we will now define *sections* of pipes to be the short (20'-30') lengths of pipe that are used to physically construct a pipeline. A *segment* is defined to be a length of pipe (perhaps many sections) of constant properties of diameter, roughness and annualized cost. *Links* are defined to be collections of segments between two nodes, the lengths of which add up to the required length of the pipeline between the nodes.

1.2 Practical Considerations

Some insights into the situation surrounding the solving of WDS problems were gained from attending an international water conference titled *Integrated Computer Applications for Water Supply and Distribution*, September 7-9, 1993. This conference is held every six years at De Montfort University, Leicester, UK. The insights derived are detailed below along with their expected impact on WDS modeling efforts.

Categories of WDS Researchers. There appear to be three distinct groups of people in the WDS business with interest in mathematical modeling:

(i) The academic researcher who may not have any practical water engineering experience, but has a variety of sophisticated theoretical and practical mathematical modeling tools available. Typically, the problems addressed are "sanitized," and it is assumed that information and data gathered are complete, and that design goals are simple, well defined and accessibly documented. Researchers will typically publish their results to get the widest possible feedback and dissemination.

(ii) The applied water engineer responsible for the operation, maintenance, and expansion of existing networks. The problems addressed are complex and messy, having uncertain existing conditions, data, goals, and funding. The applied engineer is usually responsible for water districts over geographic areas ranging from small sections of towns to large portions of a whole country (in one case all of western Australia!). Results take the form of practical solutions to day-to-day or longer term problems involving the distribution of clean, healthy water to the consumer. These results are occasionally published.

(iii) The professional consultant who caters to the water municipalities with computer software, consulting, design, troubleshooting, and other services. The problems addressed have to be practical since they need to deal with actual real world problems in order to survive in the business. The consultant may or may not have practical experience in the field but uses all the tools available including the theoretical developments, complex optimization and simulation models, and heuristics that have been proven to "work" in the past. The consultant appears to be a primary bridge between academia and the practicing water engineer. Higher level results may be published, but details of models are necessarily kept secret due to the competitive marketplace. Thus, the consultant's efforts benefit the practicing engineer, but are not in general available to the academic researcher.

There are many gray areas in, and exceptions to, the above delineations, but they serve to give academic researchers an idea of how they fit into the water systems analysis environment, thereby giving insight into how they can be most helpful in improving future water distribution systems. It is not foreseeable in the near future that many of the practical problems with messy incomplete data will be solved by academic research. When (and if) this occurs, some type of artificial intelligence and/or fuzzy logic may be required to completely analyze these situations. But in the meantime, it will be the role of the academic researcher to glean the important issues from the current situations and improve on models and algorithms by incorporating more and more of these practical considerations that face the practicing engineers and consultants. In this manner, their work will be more easily incorporated by the practitioners in their effort to reduce the cost and improve the quality of water distribution around the world.

<u>Data Centered versus Model Centered World View</u>. Apparently, many models are developed with the thought that this is the only model in existence, that a complete data base will be "plugged into" the model and everything will follow suit. This limiting "world view" is pictured in Figure 1.1 below.

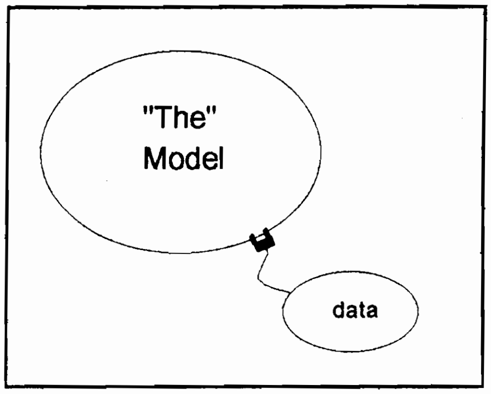

Figure 1.1 Model Centered World View

A more useful "world view" is to consider the data base as the central component, with a whole host of interconnecting models all plugged into the data base as depicted in Figure 1.2 below. To support this notion, consider that many water engineers view data collection as their most important and difficult job. From this perspective, models will come and go, but the evolving data base representing the system will always remain.

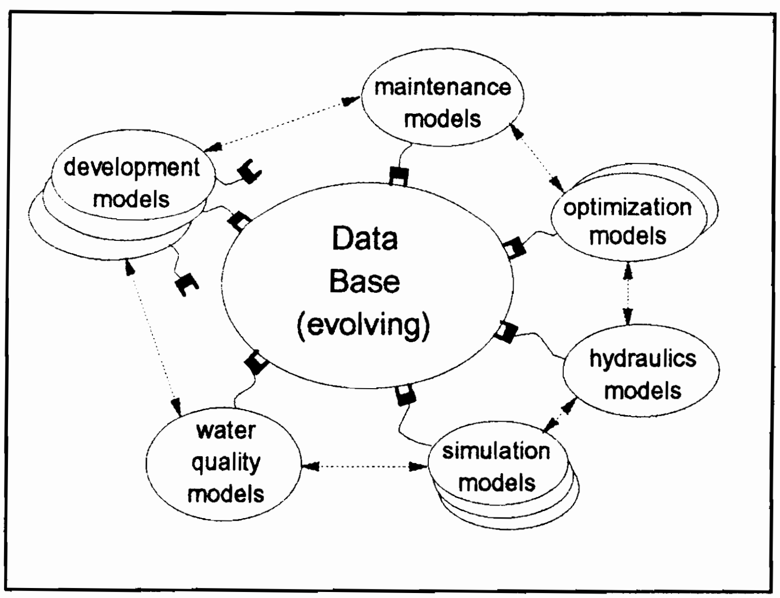

Figure 1.2 Data Centered World View

Solving Real-World Problems. As stated above, most academic research oriented problems have been simplified and sanitized from their real-world counterparts, it being assumed that all data is complete and that design goals are well defined. In reality, problems are complex and messy, having many operating conditions, a continuum of demand situations, and uncertain and unknown data, goals and funding. For example, some districts in the UK do not even know where older existing pipes are located. Other districts have as much as 40% leakage from their WDS, with projected increases in this percentage! One district in the UK develops plans assuming knowledge of only 30% of the system data at any point in time (pipe coefficients, pressures flows, etc.). In fact, a heavily discussed problem at this conference was finding the best method of locating the actual customer demand points based on billing record and postal codes (not a simple problem in London). On the other end of the spectrum, speakers from Holland and from some districts in the UK appeared to have all of the required data and were able to accurately model and predict system behavior.

Many practicing water engineers realize the benefits of the optimization process in finding lower cost solutions to WDS problem, even though the analysis and modeling may be complex. However, there are still those (probably the majority) who will not use a model that they do not understand. As a result, some designers prefer to use enumeration models that compare a finite collection of potential configurations, the primary benefit being that each candidate set of system components can be modeled explicitly using existing tools that are well understood. Unfortunately, oversimplifications occur, such as requiring that all links have a constant diameter across their whole length, and these limit the selection of solutions and thereby virtually guarantee that inefficient solutions will be chosen. Moreover, the problem of excessive computations for conducting such an enumeration still remains, even after efforts are made to trim the list of candidate solutions. Many designers do not use optimization methods at all, but prefer to use experience in hand-designing systems with the assistance of network simulators. In any case, the total design process, including such factors as reservoir design, tank fill rates, water quality, operations, data collection methods, budgeting, pipe sizing, and redundancy, is so very complex that in the near future, optimization methods will only be another tool to help the experienced water engineer in the continuing effort to keep up with the maintenance and expansion of the WDS.

There is a perception among practicing water engineers that academic researchers are too quick to claim an "optimal" answer to a problem. Not only are there many factors that are not quantifiable or accurately obtainable, but several so-called "optimum solutions" are often disproved later by better solutions obtained via new algorithms! The most well known example is the New York City water tunnel problem where the "optimum" minimum cost solution decreased from $73M to $64M to $41M (over a period of 15 years). Therefore it is important to understand that engineering judgment cannot be removed from the design process, and that unless one has proved optimality, it is best not to claim an "optimal" solution when only the best known solution is at hand.

Nowhere does there appear to be a consensus on how to define reliability in WDS's or how to incorporate it into modelling and design efforts. In fact, a major criticism of current design approaches is the tendency to attach surrogate reliability measures that do not reflect the true redundancy requirements of real systems. The perception is that either optimization models are being used to create inexpensive but unrealistic tree based designs, or that purely heuristic man-in-the-loop methods are used to create highly redundant and balanced, but expensive, designs.

In accumulating the comments and questions from many participants at this conference, the following features appear to be important in giving credibility to a model, and its accompanying implementation:

- "makes sense" (not too abstract)
- hydraulically correct
- mathematically stable (always returns with a solution)
- bug free
- good error messages
- incorporates existing pipes and commercially available replacements
- explicitly models all pipe mains
- realistic looking balanced looped networks
- self-starting without requiring "advanced" start solutions
- able to handle 10's of pumps, 100's of nodes and pipes, multiple demand situations, and messy incomplete data

It is in light of the above considerations that the following methodology is proposed.

1.3 Proposed Approach

Traditionally, pipe-reliability-and-cost models and network-optimization models are presented separately in the literature. For each type of model, inputs that are required from the other type are assumed to be known. We propose an approach for designing and generating an expansion plan for a pipe network WDS that integrates these two models into a single comprehensive design approach. There are two submodels in this approach. The first submodel is a pipe reliability and cost model that analyzes the existing pipes in the network and identifies candidate pipe segments for replacement. The second submodel in this approach is a pipe network design model that provides optimal least cost pipe sizing decisions for a fixed network demand pattern. The integrated design approach derives inputs for each submodel by conducting a preprocessing run of the other submodel, and continues this feedback process until a stable network design is obtained.

Pipe Reliability and Cost Submodel. An optimal least cost network design and replacement model requires a comparison of estimated costs of retaining existing pipes, or replacing them with new pipes, along with a consideration of the various replacement options. These determinations require the computation of optimal lifetimes for both existing pipes and new pipes, that, in turn, require estimating individual pipe segment reliabilities. Several analytical methods have been proposed by researchers that can predict the above factors under a wide range of conditions. Accordingly, this submodel will use existing statistical methodologies based on historical records of pipe breaks to estimate individual pipe segment reliabilities in order to predict future annualized costs. As a side result, failure prone, undercapacitated, or high cost existing pipe segments will be identified for removal or replacement.

Pipe Network Design Submodel. The second submodel is a comprehensive pipe network design model that provides replacement and new construction decisions for a fixed network configuration with a number of demand patterns, including the peak demand and various firefighting demand patterns. (These demand patterns specify the flow rates and hydraulic pressure levels required at each demand node.) Furthermore, this submodel includes in the network design a level of hydraulic redundancy which ensures that demand can continue to be satisfied when any one pipe is removed from the network due to a catastrophic failure or maintenance activity.

The core driver in such an approach is a single-stage network optimization process that determines a least cost pipe sizing (diameters and lengths) and energy requirements (source height) for a fixed demand pattern. The problem formulated is a nonlinear program to minimize the cost of designing pipes and energy sources subject to satisfying hydraulic flow and pressure requirements of a network. Pipe links are limited to be composed of existing pipe segments in the network and/or new pipe segments selected from commercially available pipe diameters. Existing pipe segments may be retained intact, or may be replaced either partially or completely. The pipe links are sized by varying the lengths of the component segments, each having a fixed diameter, roughness and cost. The problem formulated is a hard nonconvex optimization problem that has many local optima, different from a globally minimum cost design, and has hence proven to be difficult to solve. Researchers have studied this problem for three decades, proposing solution methodologies that yield better and better approximate solutions for various test problems. Two strategies for approaching this problem are proposed: a decomposition strategy for finding good quality local optimal solutions, and a revolutionary global optimization strategy which uses the new Reformulation-Linearization Technique (RLT) developed by Sherali and Tuncbilek (1992, 1994) (also, see Sherali and Adams, 1990, 1994).

Robustness and redundancy is accomplished in the pipe network design submodel by sequentially calling the single-stage network optimization process for each new combination of demand pattern and network configuration resulting from pipe removal. As the network design develops, the minimum pipe diameters for each link are retained, thereby building up a hydraulically redundant network that meets all the requirements for pressure and flow over the proposed demand patterns and pipe failure scenarios.

The remainder of this dissertation is organized as follows. Chapter 2 reviews the critical literature that supports this research effort. Chapter 3 describes the two submodel problem formulations and the integrated overall approach. Chapter 4 presents the theoretical algorithmic elements and practical implementation aspects for the proposed decomposition approach. Chapter 5 discussed a global optimization approach based on the Reformulation-Linearization Technique as applied to this problem. Preliminary computational results are presented in Chapter 6, and a summary of the principal contributions and conclusions, along with suggestions for future research are given in Chapter 7.

2. LITERATURE REVIEW

2.1 Pipe Reliability and Cost Models

Water lines that are constructed and installed properly in low stress areas can easily last over one hundred years. In some cities where case studies have been performed, many pipes were found to be very reliable with little or no maintenance history. However, several water mains and smaller pipes in use today were manufactured with many defects, were poorly installed, have already aged considerably, or have high stress environments. These pipes can require frequent and expensive maintenance, sometimes at an early age. When a pipe can no longer be maintained at an annual cost less than that of replacement, the pipe should be considered for replacement. However, other more subjective factors need to be considered as well, such as availability of funds and labor, convenience to the public, and safety. The many factors (alone or in combination) that contribute to water main breakage and the resulting cost of maintenance can be grouped into the following categories.

Breakage Factors:

(1) pipe factors - quality of manufacturing, size, age, material

(2) environmental factors - soil corrosiveness, water quality in the pipe, frost, soil heaving, external loads from development

(3) installation factors - quality of workmanship in laying the pipe and installing connectors

(4) service condition factors - hydraulic pressure, hydraulic gradient, water hammer

Maintenance Factors:

(1) pipe factors - size, material

(2) environmental factors - location, depth, soil type

(3) type of break - corrosion, surge, loading

Several predictive models based on historical data on many of the above factors have been developed to study pipe breakages in order to determine an optimal replacement schedule for a water distribution system.

Stacha (1978) presents a simple model for water main replacement based on the premise that a water main should be replaced when the annual cost of maintenance is more than that of replacing the main, but replacement is considered for the current year only, with no recommendations for future replacement. A segment of pipe is considered only when the frequency of repairs show a consistent increasing pattern. Also, Stacha emphasizes the intangible factors that may override the analytic financial analysis, such as water quality, safety, convenience to the travelling public, and continuous water service to the customer.

Shamir and Howard (1979) develop a model for determining the optimal year for individual pipe replacement by considering the present value dollar costs of replacement and maintenance. Replacement costs are assumed to be decreasing with time since the payout will not occur until later, and maintenance costs are assumed to be increasing since more repairs per year will be required as the failure rate increases for the aging pipe. Thus adding these cost curves as shown conceptually in Figure 2.1 gives a convex curve from which the least cost and optimal year for replacement can be found. The model requires a projected profile for the break-rate of the pipe, as well as maintenance and replacement cost trends. The authors found that an exponential equation fits the break-rate growth of their particular example well. As the model also uses the expected break-rate for the new pipe, several alternatives may need to be considered for each pipe.

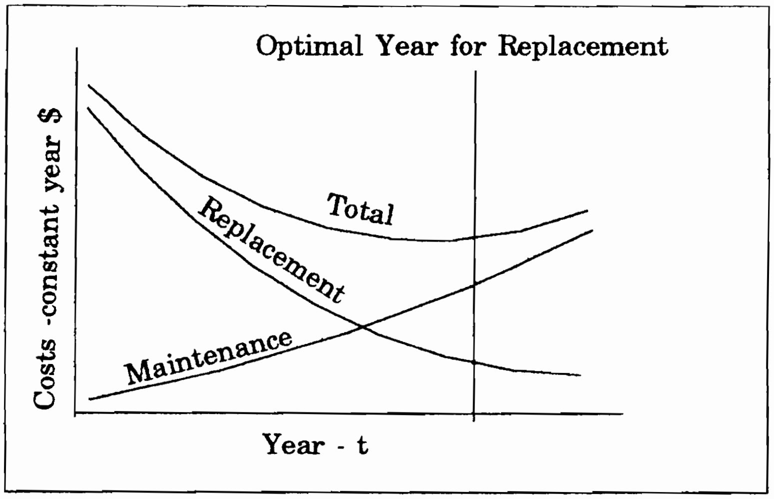

Figure 2.1 Conceptual Cost Curves

However, the authors found that for a representative example, the consideration of maintenance costs for the new pipe made little difference in the optimal year of replacement (example: 22 versus 21 yrs). As with Stacha, the authors emphasize that the model should only be used as a guide for when to replace a pipeline; practical considerations such as the availability of funds, workers, and the impact of service must be made which might outweigh financial considerations.

In two separate case studies, Clark, et al. (1982) investigate replacement cost and frequency data related to a water distribution system for a large midwestern city utility and a nearby smaller rural utility. As the definition of a "break" in a water main is difficult to determine (versus leaks, reduced capacity, etc.), the authors use repairs or maintenance events rather than only actual pipe ruptures. Data was collected on many factors for each repair, including pipe size and type, pressure, traffic over pipe, soil type, freeze and thaw information, and land development. A minority of the pipes were found to cause the majority of failures, and once a pipe started requiring maintenance, the rate of maintenance required increased over time. The expected time before the next repair was found to be also closely related to the age of the pipe. Two statistical regression equations were developed to predict/fit the probability of failure for a given number of years, based on the above factors and the number of maintenance events that had already occurred. The first equation is for pipes having no repair history, and the second is for pipes having at least one repair. A financial analysis is then performed in accordance with Shamir and Howard (1979) to determine an optimal replacement year for each length of pipe, using regression equations to determine several coefficients required in the Shamir and Howard break-rate equations. The following significant relationships can be drawn from the regression equations:

1. Metal pipes, on average, last longer before first maintenance than reinforced concrete pipes, but metal pipes accumulate more maintenance over time.
2. Larger diameter pipes last longer before first maintenance than smaller pipes
3. A high percentage of pipe overlain by residential development decreases the time to first maintenance and increases the rate of maintenance.

The authors explicitly state that general prediction of pipe failure may be possible, but precise application to a particular pipe (failure prediction) is not recommended. Rather, an application of these methods is most useful to determine some of the factors that affect break/maintenance rate.

Kettler and Goulter (1985) investigated historical pipe breakage data from a number of cities and develop statistical regression model relationships between certain factors. For example, the failure rate for pipes was determined to be strongly linked with the diameter of the pipe in roughly a linear fashion, having a decreasing failure rate with an increase in diameter. Also, failure rates for certain failure modes were determined to be increasing with time or with the number of previous breaks, while for other modes, these rates were constant over time. The authors emphasize the need for water distribution system designers to consider these types of relationships in order to determine an optimal balance between the initial cost of a system and the future performance and maintenance costs.

Andreou, et al. (1987a,b) propose a statistical methodology for analyzing break records. Nonparametric methods are used that obviate the need to hypothesize distributions beforehand. The model developed should be useful for (a) determining the future cost of various replace/repair strategies, (b) determining an optimal pipeline replacement time, and (c) estimating the network reliability. Andreou et al. (1987b) present two case studies used to develop the model discussed above. The key to the authors' method is the identification of various stages in a pipe's deterioration that are important in modelling the pipe's future breaks. For example, a slow-breaking stage (several years between failures), and a fast-breaking stage (more than one break every two years) are identified, with the transition from one stage to the next occurring after the third break in the pipe. In the early stage, the modelling emphasis is on the probability of failure and on their assessment of reliability, and not on an economic repair/replace analysis. A proportional hazards model is used to predict the failure rate, and subsequently the survival probability, of the pipe over time. The model utilizes an underlying baseline hazard rate function that represents an "aging" process inherent in the pipe, independent of added stresses. The hazard rate function need not be specified beforehand. Multiplied by this is an exponential stress factor based on a vector of factors and coefficients determined for the particular pipe or section of pipes considered. In (1987b) a "bathtub" shaped curve was found to be appropriate for the baseline hazard rate function, characteristic of a high failure rate early and late in the life of a pipe, but a low failure rate in the middle. Similar patterns are found in other types of systems that experience infant mortality and later wear out (Kapur and Lamberson, 1977). During the later fast-breaking stage, the repair/replace decision and reliability evaluation take on equal importance. The authors recommend a simple Poisson model with exponentially increasing break-rate in this stage, although the case studies in (1987b) did not show an increasing break-rate after the third break. The case study found the following factors to be most significant in analyzing breaks: pressure in the pipe, number of previous breaks, age at the time of the second break, installation period (more recent pipes worse), and amount of land development over the pipe, and pipe length.

In addition, the authors challenge two commonly reported ideas: (1) break-rate is proportional to the length of the pipe, and (2) the number of breaks increases exponentially with time once a pipe starts breaking. In contrast, the authors found that the break-rate is proportional to the square root of the length of the pipe, and that the break-rate did not increase after the third break.

Karaa, et al. (1987) present a linear programming formulation for a water distribution system resource allocation problem. Pipes having similar maintenance histories are grouped into bundles, and the decision variables are chosen to be the percentage of these bundles that is chosen for replacement. Economic inefficiencies due to repair costs, social costs and reduced hydraulic head are minimized subject to resource constraints, implemented by a capital budgeting model across a twenty year time horizon.

Li and Haimes (1992a) develop a semi-Markov process based linear programming method for determining an optimal repair/replace time for a pipe segment. The Markov states represented are the state of operation (operating or under repair) and the number of failures that have occurred. The methods of Andreou, et al. (1987a) are used to identify two stages of deterioration and their accompanying hazard rates. The formulas of Walski and Pelliccia (1982) are used to estimate the time required to repair a pipe segment failure and to estimate the accompanying cost of repair or replacement. The steady state probabilities from the Markov model are used as parameters in the linear optimization problem, which is formulated with a multiobjective function, in that the availability of the pipe is maximized while the expected cost is included as a constraint. Various noninferior levels of availability obtained at different cost levels are investigated in order to arrive at a prescribed solution.

In Li and Haimes (1992b), the authors extend their individual pipe segment model in a multilevel decomposition approach to analyze the tradeoff between system level availability and cost by solving the posed nonlinear problem by a sequence of smaller linear subproblems. The system level availability is assumed to be a weighted average of the individual pipe availabilities determined in Li and Haimes (1992a), where the weights are determined by the relative importance of each demand node. Li and Haimes do not include any hydraulic analysis of the pipes or of the network. The Alperovits and Shamir (1977) test problem is analyzed, assuming some arbitrary hazard functions, repair times and repair costs. The model outcome is a best replacement/repair resource allocation strategy in the form of a priority ordering that prescribes which pipes to assign the largest share of funding.

All the above models use annualized costs to compute the current today's dollar cost of making repairs or replacing a pipe. The importance of this topic cannot be overstated when determining optimal pipe replacement schemes. Shamir and Howard (1979) emphasize the sensitivity of this type of model to the discount rate assumed during inflation.

Hanke, et al. (1975) addressed this situation by stating that project evaluations during inflation are usually mishandled by improperly considering the effect inflation has on prices and interest rates. In particular, they note that the Water Resources Council has adopted the incorrect and inconsistent approach of projecting real prices with inflation and discounting real benefits and costs with the nominal interest rate. To correct the situation, the authors recommend that real prices (not inflated) and real interest rates (adjusted for market rate and inflation rate) be used in projecting and discounting cash flows, with an average real interest rate reported to be approximately 10.4%. They conclude that prices and interest rates must be considered simultaneously and consistently.

2.2 Pipe Network Analysis Models

The *classic pipe network analysis problem* (Collins, et al., 1978) is to solve for the network link flows and nodal pressures (or heads) when the following inputs are known: the network architecture (placement of nodes and links), supply and demand flow rates, and the pipe segment lengths and diameters (a link may be composed of several segments, each having a different diameter and/or roughness coefficient). This is, typically, a subproblem for the standard WADS design problem, in which a desirable set of pipe segment lengths and diameters also need to be determined to, say, minimize costs. There are three classical (nonoptimization) methods for solving the classic pipe network analysis problem: Hardy Cross, Newton-Raphson, and linearization. These methods all were derived from Hardy Cross' (1936) pioneering paper in 1936. Each method starts with an initial feasible solution and seeks to obtain a stationary point to the system of nonlinear equations generated from the hydraulic constraints. Revised solutions are determined by each approach using an appropriately derived set of linear equations. Convergence is guaranteed if the initial solution is assumed to be close enough to the true solution.

Each method assumes the availability of an initial flow distribution that is feasible to the flow conservation equations. In the linearization method (see Wood, 1972, Collins, et al., 1978 and Walski, 1984), the nonlinear hydraulic loop head equations are linearized during each iteration by separating out a linear component from each nonlinear head variable, and treating the remaining nonlinear term as a fixed constant for the current iteration. The resulting system of linear equations is solved by some stable variant of the Gaussian elimination algorithm. The pressure heads produced by the linear system are then substituted back into the remaining nonlinear terms to check for convergence, failing which, a new flow distribution is computed for the next iteration. The Newton-Raphson method for solving the classic pipe network analysis problem (see Fair, et al., 1966, Jacoby, 1968 and Rowell, 1979) uses the same basic approach, except that the first two terms of the Taylor series expansion of the system are used to incorporate derivative-based information into the correction terms for the heads. This leads to a system of nonlinear equations that are solved by traditional Newton-Raphson methods. The Hardy Cross method (see Adams, 1961, Eggener and Polkowski, 1976 and Rowell, 1979) applies the Newton-Raphson method, but to only one loop equation at a time.

Collins, et al. (1978) in 1978 introduced a revolutionary method for solving the classic pipe network analysis problem by formulating a convex nonlinear network flow optimization problem, which at optimality, produces flows and nodal pressures that are feasible to the given conditions. The method uses a carefully designed objective function, along with the standard balance of flow equation constraints, which have the property that at optimality, the resulting flows and nodal pressures will be feasible to the given conditions. Standard techniques such as the Frank-Wolfe and Convex Simplex methods are used to solve the optimization problem, while taking advantage of the network structure to rapidly accelerate the solution process. Limitations include the assumptions that the energy levels of the sources are fixed and that flow directions are known *a priori*. Convergence was determined to be erratic depending on the initial solution, possibly even converging to incorrect solutions. However, they were able to solve a large problem based on the Dallas, Texas network (452 nodes, including 21 reservoirs, 516 links, and 14 pumps) to within 1.3% of the best known solution.

2.3 Single Demand Pipe Network Design Models

Unless otherwise stated, in this section, we will be considering the design of looped pipe networks under a single fixed demand situation. The problem is to find the least cost combination of pipe segment lengths and diameters, and energy levels (reservoir height and/or pump settings), that satisfy the nodal pressure and flow requirements. As in the classic pipe network analysis problem, the network architecture is assumed to be known.

The traditional method for designing looped water distribution pipe networks is to assume an initial set of candidate pipe diameters and reservoir levels, solve the classic pipe network analysis problem by one of the methods cited above, in order to compute the resulting nodal pressures and link flows. If the resulting pressure is too low at some nodes, then appropriate pipe diameters are increased, and/or reservoir heads are raised in a heuristic fashion, hopefully in a cost effective fashion. This is typically done in a manual fashion. Likewise, if the resulting pressure is too high at some nodes, or higher than necessary, then pipe diameters are decreased, and/or reservoir heads are lowered. Then, the classic pipe network analysis problem is resolved with the new pipe sizes and energy levels. This process is repeated until the required pressures are met within some reasonable limits. This method depends highly on the experience of the engineer making the adjustments in pipe sizes and reservoir heights. The benefits of this approach include its intuitive nature to the water engineer designing the system, and his/her resulting familiarity with the system during the design. However, it has been shown repeatedly that computer optimization methods can regularly find better solutions at significantly lower costs.

Alperovits and Shamir (1977) in 1977 developed a successive linear programming gradient (LPG) method for designing a looped WDS. The decision variables in their model are the lengths of pipe segments of selected candidate diameters, reservoir elevations, pump capacities and operation levels, and the installation of valves. The LPG method can handle different loading conditions in the same stage of design. Hydraulic equations are formulated for each loop in the network, and for each path from a source to a demand node. The method fixes the network flows at each iteration, thereby linearizing the head loss equations in the resulting subproblem. (Note that the costs are linearized through the process of segmenting each link using different discrete pipe diameters.) The length of each pipe segment corresponding to a fixed pipe diameter is then determined by solving the resulting linear program (LP). (This implies a set of hydraulic heads.) The dual variables from the optimal LP solution are then used to determine an improving direction for the flow variables. The flow variables are changed by perturbing the solution along this direction using some heuristic rules. These heuristics require a fine-tuning for efficient use. However, good-advanced solutions can generally be found without much experience. Also, initial feasible flows must be predetermined, and the method is known to be very sensitive to the initial flow distribution. Noticeably missing from this approach, however, is the requirement to solve any classic pipe network analysis problems using Hardy Cross type methods. Alperovits and Shamir proposed and solved a small illustrative problem, that has turned out to be a challenging standard test problem (see Quindry, et al., 1979, Fujiwara, et al., 1987, and Sherali and Smith, 1993). They also solved a realistic problem representing one particular section of a city network having 52 nodes and 65 links. Candidate pipe link diameters were limited to a few alternatives in order to reduce the large number of pipe-sizing variables.

Alperovits and Shamir also state, without proof, that optimal networks will have links that contain at most two segments such that the diameters of their segments are adjacent on the list of available candidate diameters. Rowell (1979) proves this result under the assumption of strict convexity in the cost coefficients as a function of the diameters. Fujiwara and Dey (1987) further discuss this property under varying conditions, and the result was also determined to be true for multiple-loading situations.

In 1979, Quindry, et al. (1979) show that Alperovits and Shamir (1977) had missed certain terms in the gradient expression used at each iteration. They modified the LPG method to incorporate these terms and showed an improvement in solving the small test problem using the same starting solution. However, these terms significantly increase the complexity of the problem formulation, making it all but impossible to solve large general problems.

Fujiwara, et al. (1987) further modified the LPG method by deriving a complete expression for the gradient terms assuming a single source network. The procedure also incorporates a quasi-Newton direction method along with suitable line search techniques to determine step sizes, significantly increasing the computational efficiency of the algorithm. Fujiwara reiterates the need for a good starting solution, and recommends that this modified LPG method be started from the best set of flows found by some suitable heuristic method. He was able to further improve on Alperovits and Shamir's small test problem solution, but again, the task of formulating the problem for a general instance appears to be elusive, due to the large number of loop and loop interaction terms involved. Kessler and Shamir (1989) also make further improvements on Alperovits and Shamir's (1977) LPG method using projected gradient method.

Bhave (1978) formulated the single stage network optimization problem as a nonlinear problem with continuous pipe diameters for each pipe link. The problem is solved via manual manipulations in an iterative fashion. Primary links are identified in a basis tree by heuristic critical path concepts. At each iteration, the nodal heads are assumed fixed and the corresponding flows on the primary and secondary links are determined. The heads are adjusted between iterations based on a cost-head-loss ratio criterion in a fashion similar to the Hardy Cross method discussed above. The procedure is terminated when the changes in nodal heads or the improvement in the total system cost between iterations is sufficiently small. After optimization, Bhave replaces pipes having nonstandard diameters with two segments of the next larger and smaller commercially available diameters such that the combined head loss remains the same across the link. This method was applied to the Bandung, Indonesia water supply system, that has one source, 58 nodes and 90 links, reducing the total cost after only four iterations to 10% below the best currently available computer methods.

Quindry, et al. (1981), in 1981, also formulated a linear program similar to that of Bhave (1978), but the authors derived more rigorous formulas for changing the heads between iterations by taking gradients from the empirical Hazen-Williams flow equations. Unlike Bhave, the objective function is assumed to be linear. Expansion of an existing network is incorporated by including existing pipes into the linear program having zero cost coefficients in the objective function, and having an upper bound on the pipe diameters determined by the existing diameter. The authors state that no method of minimum cost optimal design of looped water distribution systems is completely satisfactory because of the vague nature of the redundancy requirements.

In a discussion of Quindry et al. (1981), Templeman (1982) highlights the difficulties and errors associated with linearizing nonlinear cost functions and further, asserts that the optimization of looped distribution networks using discrete diameter pipes is an NP-hard problem, thereby requiring solution times that are at best an exponential function of the number of pipes in the network. He concludes that discrete pipe optimization algorithms should therefore not attempt to obtain rigorous solutions, but rather, should aim at achieving "quick and dirty" heuristic solutions that approximate the optimum.

In 1983, Bhave (1983a) developed a two stage linear programming approach for optimizing multiple source looped WDSs. In the first stage, he heuristically constructs a distribution graph of primary links (core tree) using shortest path routes from source nodes to demand nodes, in order to determine the set of links to be included in the tree. These links are sized by solving a linear transportation problem, assuming the availability of continuous pipe diameters. In the second stage, some flows are assigned to the nonprimary links, and the flows in the primary links are thereby determined. Once all the flows are determined, the resulting LP problem is solved allowing, this time, only commercially available diameter segments. In a second paper, Bhave (1983b) solves a small test problem having 10 nodes (including 2 sources) and 13 links.

Jeppson (1985) introduces a practical procedure for improving upon an existing network design by identifying two dominant sources, connecting these together by a dominant path, then connecting all sources to the dominant path by the shortest paths available, and then connecting all remaining nodes to these paths by their shortest available path, resulting in a branched tree network structure. Remaining links in the network are assigned some minimum diameter for redundancy purposes. The flows and heads are computed in a heuristic manner that approximates an optimal solution. The pipe diameters are then determined by appropriate hydraulic equations, and are rounded off to the next larger or smaller commercially available diameter based on the preferences of the user. Two sample network problems are solved, yielding resulting reductions in annual operating cost of 34% and 49%.

Walski (1984, 1985) summarizes the pipe network optimization research to date, and predicts that they would become everyday tools for the water distribution system designer within a decade. He comments on the difficult, if not impossible, task of optimizing WDSs and notes that none of the existing methodologies can solve real world optimization problems. Limitations cited include a lack of ability to solve for multiple demand patterns, discrete pipe sizes, optimal design of storage reservoirs in the system, and pipe and network reliability considerations. He also emphasizes the interrelationships between pipe sizing decisions and pump, valve, and tank design and operation decisions. He concludes that optimization is helping to reduce the costs associated with typical oversized designs, but that most models that look attractive in the literature "cough and sputter when fueled with real data."

Lansey and Mays (1985) simplify the solution of the nonlinear Single-Stage Pipe Network Design Models by incorporating a network simulator into the optimization model to solve for a set of state (dependent) variables based on the value of the control (independent) variables at each iteration. General reduced-gradient optimal control theory methods are used to solve the simplified nonlinear programming problem. To assist in the solution of these subproblems, the nonlinear hydraulic constraint equations are incorporated into the objective function in an augmented Lagrangian penalty approach. The methodology as described in this paper is very general and can incorporate any number of nonlinear programming methods and network simulation models.

Both Gessler (1985) and Loubser and Gessler (1993) use enumeration to optimally design water distribution networks. The basic technique is to sequentially investigate all possible combinations of pipe diameters, pipe maintenance options, pump sizes, etc., testing each for feasibility and cost. Optimization via enumerative techniques has appeal in that this is readily understandable by applied water engineers, and that existing network analysis tools can readily be incorporated in such an approach. Unfortunately, pipe diameters are considered to be fixed across a whole link at commercially available sizes, in contrast with permitting the link to be split into segments having differing diameters. This limits the designer to only a selected set of candidate solutions, and most probably, a set not containing an optimal solution. However, they emphasize that traditional optimization techniques typically only yield a locally superior solution, and they often overlook significantly better solutions. Furthermore, they argue that enumeration techniques naturally lead to a list of Pareto optimal solutions, for which no size combination can be found that will provide higher pressure at a lesser cost. These solutions can be compared with the budget and growing demand pattern to better help the designer trade-off cost versus performance. Since the optimization process is so complex, Gessler states that optimization is only a first step in the design process toward a final solution. The combinatorial explosion of possible alternatives is readily admitted by the authors, but they insist that this problem can be significantly reduced by grouping together many similar classes of pipes so that there are fewer decision variables, and the fact that many infeasible or inferior (dominated) solutions can quickly be pruned from the list of candidates. They conclude that the method lacks mathematical evidence, but makes up for it in simplicity and flexibility.

Morgan and Goulter (1985) emphasize the importance of developing models that achieve true hydraulic redundancy, not just having each node connected by at least two pipes (connectivity), with hydraulic requirements only being met under normal operating conditions. Hence, in 1985, they developed a heuristic procedure for designing or expanding a looped WDS under multiple load and pipe failure combinations. The model is based on the traditional method for designing looped water pipe networks discussed above in this section. The Hardy Cross method (or any comparable method) is used to perform the underlying network analysis, and an LP is formulated to modify the pipe design. Since multiple load scenarios are considered, a large number of constraints are generated. These are heuristically reduced to a reasonable number.

Hobbs and Hepenstal (1989) state that optimization based methods are inherently biased toward providing solutions that appear to be more cost effective than they actually turn out to be when implemented. In particular, they note that the objective function and constraint coefficients are usually fixed at constant values when in fact these parameters are typically unknown and very likely to be random variables. The bias is due to the variation in parameters that are assumed to be constant in most models. Even though their focus is on water resources decision support systems involving the location and sizing of reservoirs, operating rules, and government policy, their observations apply to pipe sizing optimization programs as well. The authors recommend that Monte-Carlo techniques be used to further investigate operating conditions based on random parameters. They conclude that optimization based objective function values should be taken with a shaker full of salt!

A characteristic of all the LP methods reviewed thus far is the inability of the linear programs to remove uneconomical links from the network (the hydraulic gradient rises higher and higher as the pipe diameter is reduced, approaching infinity in the limit). Morgan and Goulter (1985) use a heuristic method to consider removing such links based on a minimum diameter criteria and a pipe weighting scheme during each iteration. If the new network resulting from the removal is more economical, then the link is removed, else, the link remains. Using this algorithm, they solved a medium-sized design problem having 20 nodes, 2 sources, and 37 links. Furthermore, for the New York city tunnel expansion problem proposed by Schaake and Lai (1969), they were able to obtain a better solution than Gessler (1982), with less computational effort.

Fujiwara and Khang (1990) present a two-phase decomposition approach in which the single-stage pipe network design problem is solved for new systems or for expanding old ones. In the first phase, a variation of Alperovits and Shamir's (1977) linear programming gradient method is used to obtain a locally optimal solution to this nonlinear problem. Then, the link losses are fixed, and the resulting concave problem is solved to restart the first phase with new initial flows, thereby obtaining a new, hopefully better, locally optimal solution. The procedure terminates when no progress results upon returning to the first phase. The method was applied to the New York city tunnel problem and solutions were found that were significantly better than those found by Schaake and Lai (1969), Bhave (1985) and Morgan and Goulter (1985).

All of the above papers claim to converge only to locally optimal solutions, assuming that they do converge. No one has yet found a method that is guaranteed to converge to a globally optimum solution, given the basic model formulation.

Eiger, et al. (1994) have recently presented a global optimization algorithm that using a master nonconvex primal problem to compute upper bounds and an inner linear dual problem to compute lower bounds. These problems are imbedded in a branch-and-bound algorithm to reduce the lower-upper bound gap to within a prescribed tolerance. Their application to a number of test problem have shown significant improvements over previous methods, and these are the first published lower bound results for this type of problem. However, the initial bounds on the flow variables are reduced considerably with the aid of a heuristic that is highly dependent on the initial flow variables chosen, thus true global optimality cannot be ensured. This approach amounts to solving the problem "globally" over a restricted localized region. We hope that our research with the RLT methods presented in Section 2.5 and Chapter 5 will build upon this work and prompt new research in global optimization strategies for solving this problem.

2.4 Integrated Pipe Network Design Models

Several of the pipe network design models reviewed in the previous section impose a minimum level of reliability by requiring that all the links in the proposed network have at least a minimum diameter, perhaps with differing minimum diameter limits for each link. Sometimes, this is accomplished directly after a core tree design is specified, as in Jeppson (1985); at other times, this is simply the result of the optimization process over the entire network, as in Alperovits and Shamir (1977). In this section, we review models that integrate reliability into the design process more completely.

Delfino (1975), in his dissertation, proposes solution methodologies for designing minimum cost WDSs. He considers a looped WDS design problem, and discusses the difficulties in formulating a reliability requirement which would ensure that in the event of a link failure, the remaining network would still be capable of supplying each node via an alternate route. Hydraulic integrity under such emergency loading conditions were not considered. Of particular interest, is Delfino's theorem that if reliability requirements are not specified, the minimum cost solution to a looped WDS design problem under a single demand pattern would be a spanning tree design. This theorem is used in much of later research in justifying the design of WDS using a core tree design. The proof assumes that a continuum of pipe diameters are available (see Rowell, 1979). However, when the pipe diameters are restricted to be selected from among standard discrete values, this result does not necessarily hold.

Next, we review Rowell's (1979) dissertation, in which he developed a three level system of models for the network design of a WDS under peak and emergency loading conditions. The objective function maximizes nodal pressures in the network when links or pumps are removed, or when fire demand loads are applied. Normal operating condition requirements are enforced via constraints. Capital and maintenance costs are incorporated as constraints subject to the available budget. In the first level of design, Rowell determines a core spanning tree subgraph corresponding to the list of available links. At the second level, redundant links are introduced for reliability purposes, and at the third level, the detailed design for the pipes and energy sources (pump settings and water tower heights) is accomplished. However, this study does not include the consideration of expanding designs for cities having existing pipe networks in place.

To design the core tree, Rowell formulates a minimum cost flow problem to find a shortest path tree, where the lengths of the network links are weighted by the necessary water flow through the tree required to satisfy the peak load demand. To connect the potential multiple trees (each attached to at least one source), he arbitrarily adds links between the trees, perhaps the shortest such connecting links. Reviewing a number of research papers, Rowell found that trees determined in this heuristic fashion were found to be consistent with the final network design prescribed in these papers. (In all cases, the core tree was a subgraph of the resulting network.)

Two methods were proposed for determining the redundant links at the second level of the design process. An easy to solve set-covering model enforced connectivity, but did not provide further design information. Alternatively, a difficult to solve (general integer program) flow-covering model enforced connectivity along maximum flow routes, and provided the engineer with a set of minimum pipe diameters for the redundant links.

Once the final network configuration was determined, a detailed design was prescribed using a modification of Alperovits and Shamir's (1977) or Quindry, et al.'s (1979) algorithms. Rowell experienced problems in finding feasible solutions, given the multiplicity of loop constraints required to solve medium sized problems (24 nodes, 33 links, 2 sources). He remedied these problems by formulating various relaxations to the loop constraints, and by reducing the large number of pipe-sizing variables, allowing only three pipe diameter candidates for each link.

In Seller's (1984) discussion of Rowell and Barnes (1982) procedure (derived from Rowell's (1979) dissertation), Seller states that standard pipe network software does not allow for contingencies arising from peak demand occurrences at demand points at different times. He asserts that the conservative assumption that peak demands across the network are simultaneous leads to inefficient designs, particularly if the resulting designs are tree-based. He concludes that loops are not added to a network merely for redundancy, as commonly assumed, but add significantly to the ability to service nonsimultaneous downstream peak demand patterns.

Loganathan and Sherali (1990) developed a two-phase network design methodology that solves a single-stage pipe network design problem having an imposed reliability constraint. The reliability constraint requires that each demand node be connected to a source node by at least two arc-disjoint paths, thereby ensuring that connectivity is retained under a wide range of failure conditions, including the failure of any one pipe in the network. Since optimization by its very nature removes redundancy from the network, and simply imposing minimum diameter pipes for each link does not truly bring about an optimal design, the authors first develop an optimization method based on Rothfarb, et al.'s (1970) approach, along with efficient network flow programming techniques (see Bazaraa, et al., 1993), in order to find a good quality core tree design. In the second phase, redundant links are selected by heuristically choosing links that satisfy the reliability requirement while minimizing the total length of new pipes. The redundant links are iteratively sized until the hydraulic upper and lower bounds are satisfied. Finally, the two phases are coordinated in a computationally tractable fashion using a heuristic technique of "guess-and-verify" to ensure that hydraulic consistency is established in the final solution and that the reliability constraints are not myopically imposed without having investigated readily available lower cost feasible solution alternatives. The algorithm was successfully applied to the single source Alperovits and Shamir (1977) test problem and to Rowell and Barnes' (1982) multisource network to demonstrate its efficiency and competitiveness. The authors conclude that the main limitation of this approach is that the procedure for coordinating the two phases is not yet sufficiently defined to allow for automated implementation as the critical decisions in this procedure are left to the judgement of a skilled user.

Fujiwara and Tung (1991) focus on the reliability issue by limiting the reduction in the total water supply due to pipe failures. They assume that links are restricted to a single pipe diameter and that continuous pipe diameters are admissible. An initial solution is obtained via a nonlinear maximum flow problem, and then reliability is improved by considering marginal reliability improvements and cost increases when pipe link failures occur. The marginal increases are taken with respect to increases in the diameter of the remaining pipe links, and this directly improves physical system reliability. Like its contemporaries, the algorithm is heuristic in nature and therefore does not necessarily lead to optimal solutions.

The papers reviewed in this chapter display a wide range of breadth and depth of analysis and complexity using many diverse methods. None however, integrates the methods of pipe failure analysis, network reliability and optimization. In the next chapter, we present the basic methodology proposed for combining these elements into an integrated structure that accomplishes the optimal replacement and design strategies for a reliable water distribution system across a number of demand patterns and load requirements.

2.5 Reformulation-Linearization Technique (RLT)

The Reformulation-Linearization Technique is a new optimization scheme whereby various classes of nonconvex problems that arise in many production planning, economics and engineering design contexts can be addressed. In some cases, the technique provides solutions that are within a tolerance of optimality (converging infinitely), and in others, globally optimal solutions can be found in finite time. At the heart of this methodology is a procedure developed for generating a hierarchy of progressively tighter, higher dimensional, linear programming or convex representations for the underlying polynomial zero-one, continuous and mixed-integer programming problems. This procedure can eventually lead to an explicit characterization of the convex hull of optimal solutions in the discrete case, and also in some continuous situations.

Glover and Wolsey (1974) and Glover (1975) help to lay the groundwork for RLT by proposing relaxations of 0-1 polynomial programs via various linearizations of nonlinear terms. These linearized constraints are designed to ensure the equivalence of the resulting mixed-integer problem and the original nonlinear 0-1 problem. The RLT, on the other hand, reformulates even 0-1 linear 0-1 mixed integer problems, providing for an automatic constraint generation scheme that produces tight linear programming relaxations. Moreover, it applies even to continuous, nonlinear polynomial programming problems.

Several recent advances have been made in the development of branch-and-cut algorithms for discrete optimization problems and in polyhedral outer-approximation methods for continuous nonconvex programming problems. (For examples, see Hoffman and Padberg, 1991, Nemhauser and Wolsey, 1988, Horst and Tuy, 1993, and Sherali, 1991.) At the heart of these approaches is a sequence of linear programming problems that drive the solutions process. The success of such algorithms is strongly tied to the strength or tightness of the linear programming representations employed.

The particular Reformulation-Linearization Technique developed by Sherali and Adams (1989, 1990) addresses mixed-integer zero-one linear programming problems. They construct various polynomial factors of degree *d* comprised of the product of some *d* binary variables and their complements, called alphabets. The factors are then used to multiply each of the constraints defining the original problem to create a nonlinear polynomial mixed-integer zero-one programming problem. Next, the relationship $x^2=x$ is used for each binary variable x, and a linear term is substituted for each nonlinear term. Relaxing integrality, the problem region is re-cast into a higher dimensional polyhedral set. As the degree d increases, it is shown that this polyhedron more and more closely approximates the convex hull of the original feasible region, and for d equal to the number of binary variables, the convex hull is actually realized. As far as using RLT as a practical computational aid is concerned, one option is to simply work with the relaxation programs at the first level. Adams and Johnson (1994), Adams and Sherali (1986, 1993), Sherali and Adams (1984), Sherali and Alameddine (1992), Sherali and Tuncbilek (1992b, 1994), and Sherali, Ramachandran and Kim (1992) have shown how efficient algorithms can be constructed for various classes of problems and applications. We refer to Sherali and Adams (1993) for a detailed summary of the these developments and the basic framework of the RLT as applied to different class of problems.

Our implementation of RLT on the pipe network design problem is based on Sherali and Tuncbilek's (1992a) work on global optimization in polynomial programming. They prescribe an RLT process that employs suitable polynomial factors to multiply constraints in order to generate additional polynomial terms and constraints, which upon linearization through variable redefinitions, produces a tight linear programming relaxation. The resulting relaxation is used in concert with a suitable partitioning scheme to develop an algorithm that is proven to converge to a global optimum. We have built upon this strategy in this work to accommodate intermediate nonlinear constraints approximations, as well as to include the inclusion of nonpolynomial nonconvex constraints that have rational versus integer exponents.

We now proceed to develop a formal mathematical formulation for our network design problem, and to prescribe an overall integrated approach to address this problem.

3. MODEL FORMULATIONS AND INTEGRATED APPROACH

3.1 Integrated Design Approach

Traditionally, pipe breakage and cost analysis models are run with the assumption that, "If a pipe is considered for replacement, then the length, diameter and type of a new pipe is known." However, in a realistic situation, these parameters are determined as a result of some hydraulic analysis, perhaps via an optimization process. Likewise, traditional optimization models presuppose the knowledge of annualized capital and maintenance costs for the various sizes and types of new pipes. But these inputs are not readily available unless a reliability analysis is performed to determine when a pipe will be replaced, how much it will cost, when maintenance events are expected to occur and how much they will cost, and when the replaced pipe is itself replaced. We address this issue in a new way by integrating the hydraulic analysis from some preliminary optimization runs into the reliability and cost analyses. These preliminary runs assume that the network is designed with all existing pipes in place and with new pipes placed in parallel to them, along with new pipes installed in the expanded part of the network. This gives us a good estimate of the hydraulic properties (flow and pressure gradient) that are desirable in each of the pipe links, and in turn, enables us to estimate the size of pipe that will be replaced in the reliability and cost analysis. Based on this, we decide if an existing pipe is a good candidate for replacement, and, after such replacement decisions have been made, the pipe network design submodel described in Section 3.2 is used to prescribe the actual pipe sizes and hydraulic properties. If necessary, the process can be repeated if the prescribed flow and pressure gradients are greatly different from the estimated ones.

This feedback integration process of the two submodels is depicted in Figure 3.1, and proceeds as follows:

I. Preprocessing Cost Analysis. First, the reliability and cost submodel is run for all commercially available diameters in order to determine an optimal life and annualized costs for the new pipes to be used in the network design optimization submodel.

II. Preprocessing Flow Analysis. Using the annualized costs for the new pipes from step I, the pipe network optimization submodel is run for a representative demand pattern with all existing pipes in place, and with the to-be-designed new pipes added where prescribed, as well as in parallel to each existing pipe along its entire length. (Note that this assumes

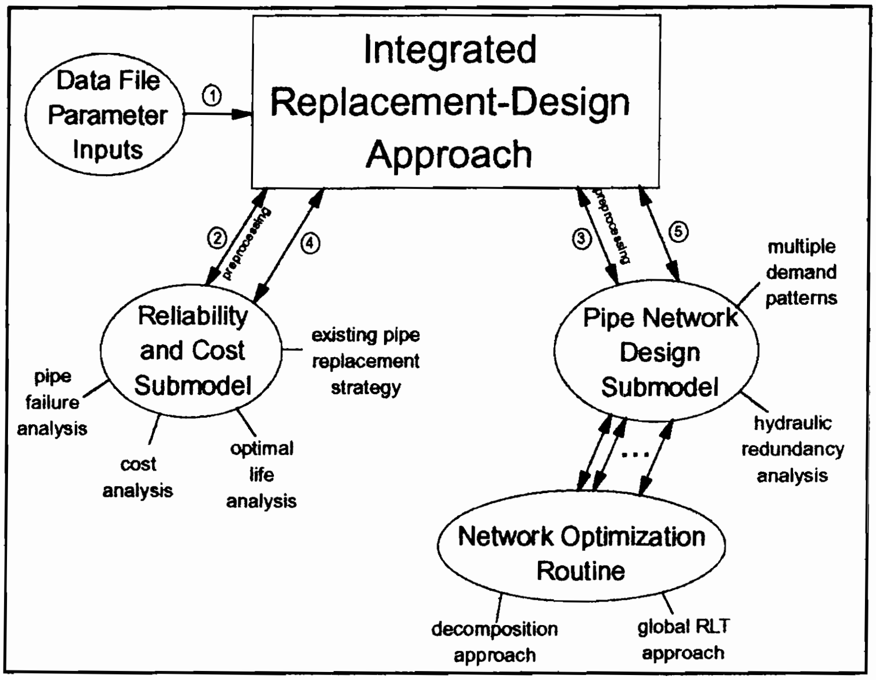

Figure 3.1 Hierarchy of the Integrated Replacement-Design Approach

that a prescribed set of connections have been designated, perhaps enforcing a connectivity based reliability. For example, the design might ensure that each demand node remains connected to a source node that can feed it whenever any single link in the system fails or is disconnected.) The resulting solution suggests a baseline flow for each link in the network that takes into account the new demands along with the existing network structure in the expansion plan. (Alternatively, the submodel can be run from "scratch", that is, with all existing pipes removed, and with new pipes in their place.)

III. <u>Pipe Reliability and Cost Submodel</u>. For each existing pipe segment, the pipe reliability and cost analysis submodel is run using the baseline flow from step II to compute the annualized expected cost over the 40 year time horizon. This cost is determined using a suggested replacement diameter that does not reduce the hydraulic gradient in the pipe, along with the optimal year of replacement. If the replacement falls within the current budgetary cycle (say, 5 years), or if the pipe segment satisfies any other criterion for replacement, the existing section is identified for replacement in the final network design. Otherwise, the existing section is retained in the final network design.

The specific repair and maintenance models assumed in this submodel are presented in Section 3.3.

IV. Pipe Network Design Submodel. The pipe network design optimization submodel is now run again, using the annualized costs computed in steps I and III, to prescribe a set of pipe section diameters and energy source levels, and a corresponding set of resulting hydraulic pressures and flow rates for each node and link.

The pipe flow rates prescribed in this step (IV) will not necessarily be the ones estimated by the baseline run of the optimization submodel in step II, wherein design improvements may result from repeating steps III and IV, using the new flows from the previous iteration of step IV as the baseline flows in step III.

V. Implementation. Prioritize the implementation of the new and replacement pipes using hydraulic property needs of the evolving system, the costs involved, budgetary limitations, and management objectives.

We will present the pipe network design submodel first, since elements in its formulation will be needed for the pipe reliability submodel.

3.2 Pipe Network Design Submodel Formulation

The pipe network design submodel is a comprehensive framework for providing replacement and new construction decisions. A fixed network configuration is assumed along with a number of demand patterns, including the peak demand and various firefighting demand requirements. These demand patterns specify the flow rates and hydraulic pressure levels required at each demand node. Furthermore, this submodel includes in the network design a level of hydraulic redundancy which ensures that these demand patterns can be met with any one pipe removed from the network due to a catastrophic failure or maintenance activity.

The core driver in such an approach is a network optimization routine that determines the least cost pipe sizing (diameters and lengths), and energy requirements (source height), for each fixed demand pattern on a fixed network configuration (perhaps with the latter being iteratively modified by removing one link at a time from the original network). The problem formulated is a nonlinear program to minimize the cost of designing pipes and providing energy sources subject to satisfying hydraulic flow and pressure requirements. Pipe links are limited to be composed of existing pipe segments in the network and/or new pipe segments selected from commercially available pipe diameters. Existing pipe segments may be retained intact or may be replaced either partially or completely. The pipe links are sized by varying the lengths of the component segments, each having a fixed diameter, roughness and cost.

Pipe costs are generally nonlinear as a function of diameter but linear as a function of length. Therefore, the cost contribution from the pipe sizing decisions modelled in the foregoing fashion is linear. The energy sources are designed by determining the heads at pre-located reservoirs. These costs are assumed to be linear over the region of interest, and thus, the objective function is linear. However, the constraints are nonlinear due to the nature of the hydraulic pressure constraints over the pipe segments. Thus the network optimization problem is a hard nonconvex optimization problem that has many local optima, different from a globally minimum cost design, and has hence proven to be difficult to solve. Researchers have studied this problem for three decades, proposing solution methodologies that yield better and better approximate solutions for various test problems. We propose two new strategies for approaching this problem:

(1) Decomposition Strategy. In the first strategy, we develop a new formulation that permits us to take advantage of special network structures inherent within this problem. The problem is decomposed into two levels of problems, a main or master problem and a series of projected subproblems. In the master problem, a set of design variables are identified, with penalty functions being used to guide the feasibility of the design. For each fixed setting of the design variables, a convex linearly constrained subproblem is formulated that can easily be solved on an augmented network. The subproblems are always feasible but may generate artificial flows in the network. The artificial flows are penalized in the main problem in a fashion so that optimal solutions with zero artificial flows are guaranteed to be optimal to the original problem as well. Conjugate gradient search techniques are used to find improving directions in the design variable space. The algorithm terminates when the projected gradient of the design variables vanishes at a locally optimum solution. Thus the hard, nonconvex problem is projected onto spaces where subproblems are easy to solve, resulting in an effective search strategy for the original problem. Such an approach makes the method suitable for handling even large-scale networks. This strategy ensures that good quality solutions can be obtained without using advanced-start solutions from other sources. However, like all its predecessors, it does not guarantee that the solution obtained is globally the best solution available.

(2) Global Optimization Strategy. Our second proposed strategy theoretically guarantees the generation of global optimal solutions, and in based on the Reformulation-Linearization Technique (RLT) developed by Sherali and Tuncbilek (1991) (also see Sherali and Tuncbilek, 1994, and Sherali and Adams, 1990, 1994). Given the structure of the pipe network design problem, a suitable RLT approach is be developed that exploits the structure of the problem and generates tight linear approximations. Such approximations or relaxations are embedded within a branch-and-bound framework to determine a global optimum to the original nonconvex problem. Since a good quality starting solution is highly beneficial in providing an initial upper bound in this global optimization strategy, we can use the results from the decomposition strategy for this purpose.

A sequential design approach is proposed for the overall problem by integrating the foregoing two submodels in the following fashion. A preliminary run of the pipe reliability submodel is used to estimate the annualized costs and to identify pipes for potential replacement. After determining these costs, the network is designed in a sequential fashion over a number of stages. First, solutions from the pipe network design submodel are obtained for the peak-load pattern, and for separate firefighting demand or other demand pattern scenarios. The maximum pipe sizes required across these demand patterns are retained for each link as lower bounds for future sizing, in order to ensure that hydraulic requirements are met under the specified conditions. This network design is called the first stage design. Next, the pipe links in the network are sequentially removed under the given load patterns (perhaps with a reduced pressure requirement such as the upper 80 percentile level), and the network is re-designed, updating the pipe sizes for each stage based on the solution obtained from the previous stages. At each step, the largest pipe sizes from the current and previous stages are retained to allow the design to be feasible to all the conditions imposed thus far. Links having the largest diameter pipe segments are removed in the initial stages to accelerate the impact on the network design, thereby simplifying later iterations in the sequence. (If the removed link disconnects part of the network from a source, then a redundant parallel link is proposed in a preprocessing step. This parallel link is then sized to run alongside the removed link.) The result is a reduced cost network design that meets all the requirements for pressure, flow and reliability over the proposed demand patterns and pipe failure scenarios.

3.2.1 Network Optimization Problem Formulation

To describe the network optimization problem, consider a region in which a set of reservoirs and key demand locations are specified. Let these be designated as the nodes of the distribution network, and be identified by an index set $N=\{1,...,n\}$, with $S \subset N$ denoting the set of sources and $D=N-S$ denoting the set of demand locations. Define the quantity b_i as the net water supply or demand rate corresponding to node $i \in N$. By convention, we will assume that $b_i \geq 0$ for $i \in S$ and that $b_i \leq 0$ for $i \in D$. As a necessary condition for feasibility, we assume that the total supply rate exceeds the total demand rate. For convenience, define a special dummy demand node 0 having demand

$$b_0 = -\sum_{i \in N} b_i \leq 0. \tag{3.1}$$

Next, let us consider the set of pipes connecting node pairs i and j in N. These pipes may either be existing ones or may be potential connections that can be installed. Each pipe can permit the flow of water in either direction, depending on the pressure head differential along the pipe. For each existing or potential pipe between some node pair i and j in N, we create a pair of directed arcs or links (i,j) and (j,i) in the network. As will be evident below, the problem solution will admit a positive flow in at most one of each such pair of links, thereby deter-mining the direction of flow in the (undirected) pipe connection. Let A denote this set of directed arcs of the network, and let Q_{ij} denote the (equilibrium) flow on link $(i,j) \in A$. For convenience in the model formulation, we will assume knowledge of Q_{ij}^U as some redundant upper bound on the flow Q_{ij} at an equilibrium solution for any network that includes the link (i,j).

For each $(i,j) \in A$, let $L_{ij}\,(=L_{ji})$ denote the pipe length corresponding to a connection between nodes i and j. We will assume that each pipe is constructed of segments comprised of lengths of standard available diameters chosen from the set $\{d_k, k=1,...,K\}$, where $d_k > 0$ for all k. (The pipe thickness is assumed to be related to the pipe diameter in a manner that can provide the required structural strength.) Additionally, we create a set of arcs (i,0), $i \in S$, in the network, in order to represent the slack left over in the supply at each source node. This artifice also facilitates the handling of the fixed head requirements.

As far as the existing links in the network are concerned, let a_{ijk} be the existing length of link $(i,j)\epsilon A$ of diameter d_k. Note that $a_{ijk} = 0$ for k=1,...,K in case there is no existing connection between nodes i and j, and that these quantities sum to L_{ij} otherwise. Furthermore, define the decision variable $0 \le X_{ijk} \le a_{ijk}$ as the length of existing link (i,j) of diameter d_k that is selected for continued use, rather than being discarded or replaced, for each $(i,j)\epsilon A$, and each d_k, k=1,...,K. (The X variables will actually be fixed by the prior run of the reliability and cost submodel as discussed in Sections 3.1, later in this section, and in Section 3.3.) Similarly, for each $(i,j)\epsilon A$ and each d_k, k=1,...,K, define the decision variable $x_{ijk} \ge 0$ as the length of link (i,j) of diameter d_k that is newly constructed for use.

For notational ease, henceforth, any symbol for which the subscripts are dropped will be assumed to represent a vector of the corresponding subscripted terms. Also, note that for each k=1,...,K, we have $a_{ijk} \equiv a_{jik}$, $X_{ijk} \equiv X_{jik}$, and $x_{ijk} \equiv x_{jik}$, for each $(i,j)\epsilon A$, and so only one of each pair (say with i< j) needs to be defined.

Next, let us define the costs associated with the links in the network. For the new links, let c_{ijk} be the annualized construction and maintenance (lining and cleaning) cost per unit length of link $(i,j)\epsilon A$ of diameter d_k, k=1,...,K. As far as the existing pipes are concerned, consider a pipe $(i,j)\epsilon A$ of unit length and diameter d_k, and assume that this pipe will last up to time T_{ijk}, after which it will be replaced by a similar or improved technology pipe. By considering the maintenance costs up to a time T_{ijk} and the annualized construction plus maintenance costs from T_{ijk} onwards, one can calculate an equivalent annualized cost component C_{ijk} per unit length applicable from the present time onwards. This is the cost coefficient which is used with the variable X_{ijk} in the model. (See Section 3.3 for the development of the cost coefficients c_{ijk} and C_{ijk}.)

Now, let us consider the energy heads at the various nodes $i\epsilon N$ in the network at an equilibrium solution. For each node $i\epsilon N$, let E_i denote the ground elevation of node i, and let H_i denote the established head above E_i. For the source nodes $i\epsilon S$, let F_i denote the maximum available energy (fixed) head. We will assume that there is an additional opportunity to raise the head at each source $i\epsilon S$ by an amount H_{si} (to be determined), at an (annualized) cost $c_{si} > 0$ per unit energy head, as suggested by Rowell and Barnes (1982). Furthermore, for each demand node $i\epsilon D$, suppose that there is a requirement that the head (H_i+E_i) at this node at flow equilibrium lies in the interval $[H_{iL},H_{iU}]$, where $H_{iL} \le H_{iU}$. If H_L^* denotes the largest lower bound H_{iL} for a demand node $i\epsilon D$, we will make the following reasonable assumption.

$$H_L{*} \leq \text{minimum } \{F_i \text{ for } i \in S,\ H_{iU} \text{ for } i \in D\} \tag{3.2}$$

The pressure loss (or head loss) in a pipe due to friction, $[(H_i+E_i)-(H_j+E_j)]$, depends on the pipe characteristics such as diameter, roughness, and length, and the water flow rate through the pipe. The frictional head loss in a segment of pipe with smooth flow can be approximately determined through the empirical Hazen-Williams equation as follows (see Alperovits and Shamir, 1977 and Walski, 1984):

$$\phi(Q, C_{HW}, d, x) = 1.52\text{x}10^4 (Q/C_{HW})^{1.852} d^{-4.87} x \tag{3.3a}$$

where

ϕ = pressure head loss in a smooth flow pipe segment (meters)

Q = water flow rate in the pipe (meters^3/hour)

C_{HW} = Hazen Williams Coefficient based on roughness and diameter (approximately 140 for a very smooth new pipe, 100-120 for an old pipe in good condition, and 40-80 for a pipe with significant tuberculation)

d = pipe diameter (centimeters)

x = pipe length (meters).

The above Hazen-Williams equation is the traditional one used in the technical literature. However, Walski (1984) states that the smooth flow conditions for which it applies are a rarity in actual systems. He recommends using the following rough flow Hazen-Williams equation:

$$\phi(Q, Q', C_{HW}, d, x) = \frac{1.52\text{x}10^4\, Q^2}{C_{HW}^{1.85}\, d^{4.87}\, (Q')^{0.15}}\, x \tag{3.3b}$$

where

ϕ = pressure head loss in a rough flow pipe segment (meters)

Q' = water flow rate at which C_{HW} was measured (meters^3/hour)

We will use equation (3.3b) assuming Q'=100, except when comparing results with the literature, in which case we will use equation (3.3a). Thus (3.3b) reduces to

$$\phi(Q, C_{HW}, d, x) = \frac{7.6x10^3\,Q^2}{C_{HW}^{1.85}\,d^{4.87}}\,x. \tag{3.3c}$$

In our model, the head loss in a pipe that has several potential segments of varying diameter and roughness can be computed as follows

$$\begin{aligned}\phi_{ij}(Q_{ij}, x_{ij\cdot}, X_{ij\cdot}) = &\sum_{k=1}^{K} \phi(Q_{ij}, Q'_{ijk}, C_{HWx_{ijk}}, d_k, x_{ijk}) \\ &+ \sum_{k=1}^{K} \phi(Q_{ij}, Q'_{ijk}, C_{HWX_{ijk}}, d_k, X_{ijk})\end{aligned} \tag{3.4}$$

where

$x_{ij\cdot} \equiv (x_{ijk},\ k=1,\dots,K)$

$X_{ij\cdot} \equiv (X_{ijk},\ k=1,\dots,K)$

$\phi_{ij}(Q_{ij}, x_{ij\cdot}, X_{ij\cdot})$ = frictional head loss in link $(i,j) \in A$

$C_{HWx_{ijk}}$, $C_{HWX_{ijk}}$ = Hazen Williams Coefficients for x_{ijk}, X_{ijk}, respectively.

The network optimization problem (<u>NOP</u>) can then be formulated as follows:

<u>NOP</u>: <u>Minimize</u>

$$\sum_{\substack{(i,j)\in A \\ i<j}} \sum_{k=1}^{K} c_{ijk} x_{ijk} + \sum_{\substack{(i,j)\in A \\ i<j}} \sum_{k=1}^{K} C_{ijk} X_{ijk} + \sum_{i\in S} c_{si} H_{si} \tag{3.5a}$$

<u>subject to</u>

$$\sum_{j:(i,j)\in A} Q_{ij} - \sum_{j:(j,i)\in A} Q_{ji} + Q_{i0} = b_i \quad \text{for each } i \in S \tag{3.5b}$$

$$\sum_{j:(i,j)\in A} Q_{ij} - \sum_{j:(j,i)\in A} Q_{ji} = b_i \quad \text{for each } i \in D \tag{3.5c}$$

$$-\sum_{i\in S} Q_{i0} = b_0 \quad \text{for node } 0 \tag{3.5d}$$

$$(H_i+E_i) - (H_j+E_j) = \begin{cases} \phi_{ij}(Q_{ij}, x_{ij\cdot}, X_{ij\cdot}) & \text{if } Q_{ij} > 0 \\ \leq 0 & \text{if } Q_{ij} = 0 \end{cases} \quad \text{for each } (i,j) \in A \tag{3.5e}$$

$$H_i+E_i \leq F_i+H_{si}, \text{ with equality if } Q_{i0} > 0, \text{ for each } i \in S \tag{3.5f}$$

$$H_{iL} \leq H_i+E_i \leq H_{iU} \text{ for each } i \in D \tag{3.5g}$$

$$H_{si} \geq 0 \text{ for each } i \in S \tag{3.5h}$$

$$\sum_{k=1}^{K}(x_{ijk}+X_{ijk}) = L_{ij} \text{ for each } (i,j) \in A,\ i<j \tag{3.5i}$$

$$0 \leq Q_{ij} \leq Q_{ij}^{U} \text{ for each } (i,j) \in A \tag{3.5j}$$

$$x_{ijk} \geq 0,\ \ 0 \leq X_{ijk} \leq a_{ijk} \text{ for each } (i,j) \in A,\ i<j, \text{ and } k=1,\ldots,K \tag{3.5k}$$

The objective function, Eq. (3.5a), in the above model denotes the total annualized construction plus operating costs. The constraints (3.5b) - (3.5d) enforce the conservation or continuity of flow at each node in the network. The constraints (3.5e) represent the conservation of energy or head loss constraints for each pipe in the direction of positive flow. Note from Eq. (3.5e) that $Q_{ij}>0$ implies that $(H_i+E_i) > (H_j+E_j)$, and so we will never have both Q_{ij} and Q_{ji} positive in any feasible solution. The constraints (3.5f) represent the maximum variable head available at each source node $i \in S$, constraints (3.5g) and (3.5h) represent minimal and maximal head levels enforced at each node, and constraints (3.5i) establish the appropriate pipe lengths for constructed pipes. Finally, constraints (3.5j) and (3.5k) represent logical restrictions.

Now, define the set of *design variables* as $w=(x,X,H_s)$ and let $W = \{w$: constraints (3.5h), (3.5i), and (3.5k) are satisfied$\}$. Recall that the X decision variables are actually determined beforehand in the reliability and cost model. This will facilitate the detailed search process as will be explained in Section 4.2. In the next chapter we will use these design variables to reformulate or decompose the model in a manner that will expose and render exploitable its special network substructure.

3.3 Pipe Reliability and Cost Submodel Formulation

The goals of this section are to formulate a viable pipe reliability and cost model and to describe its integration into the pipe network design submodel. Elements and structures of existing models will be brought together to provide the necessary inputs for the integrated WDS design approach that we propose. Each utility will, in general, have its own economic analysis model based on historical data, local trends and current policy. We formulate only one such viable alternative in the following paragraphs. The details of this model assume a hypothetical water utility, and do not involve any new on-field empirical study in this aspect of water engineering. However, the way in which existing pipe breakage and cost models are used to derive inputs for the pipe network design optimization process is crucial for understanding the integration of the two submodels.

The model formulated is used to predict the annualized costs over a 40 year time horizon for each existing pipe segment in the water distribution system network, to determine which of these will be replaced in the current planning period, and to predict the annualized cost of installing new segments of pipe at various times. In order to project pipe failure rates, Hazen-Williams Coefficients, and maintenance and replacement costs, we will use a combination of existing statistical models from Shamir and Howard (1979), Quindry et al. (1981), Walski (1984, 1987) and Ketler and Goulter (1985).

The basic pipe reliability/failure regression equation model proposed by Shamir and Howard (1979) in analyzing historical data, and recommended by Walski (1984, 1987) as a useful approximation for projecting future breaks, is

$$N(t) = N_0 \, e^{b\,(t-t_0)} \tag{3.6}$$

where

$N(t)$ = break rate in year t (breaks/year/km)

N_0 = initial break rate in year t_0 (breaks/year/km)

b = rate coefficient ($year^{-1}$)

t = time (year)

t_0 = base year.

This "break rate" function is actually called a hazard function in the reliability literature. The corresponding failure density function for our exponentially increasing break rate is the extreme value distribution. The coefficient b has the physical interpretation as the rate of change of the break rate per year. Shamir and Howard found rate coefficients to lie in the range between 0.05 and 0.15 per year and initial break rates for 1961 to lie in the range between 0.327 and 0.818 breaks/year/km. Walski (1987) reports rate coefficients in the range between 0.01 and 0.10 per year. Since break rate is a function of the breakage factors listed in Section 2.1 (pipe, environmental, installation and service condition factors) for each pipe segment having unique properties and environment, pipes can be categorized for analysis or more complex linear regression equations based on a multiplicity of factors can be developed as in Clark et al. (1982).

In a simpler model considering only the diameter as a factor, Ketler and Goulter (1985) found that failure rate decreased as diameter increased, with a strong linear tendency if the analysis was confined to a single city. For the city of Philadelphia, the failure rate for pipes between 4" and 16" was approximately

$$N(D) = .3 - .01D \tag{3.7}$$

where

N = break rate (breaks/km/year)

D = Diameter of pipe (4 to 16 inches).

The authors hypothesize that the relationship becomes nonlinear for larger pipe diameters, with decreasing slope as the diameter increase.

Since we do not have any direct water utility data, we will combine the above models using some representative parameters to form a simple failure rate model for a hypothetical utility that takes into account the dependence on time and pipe diameter. We will assume an exponential decaying relationship for pipes having larger diameters that is consistent with the slope and value of equation (3.7) at D=16" (see Figure 3.2 for a plot of initial break rate versus diameter).

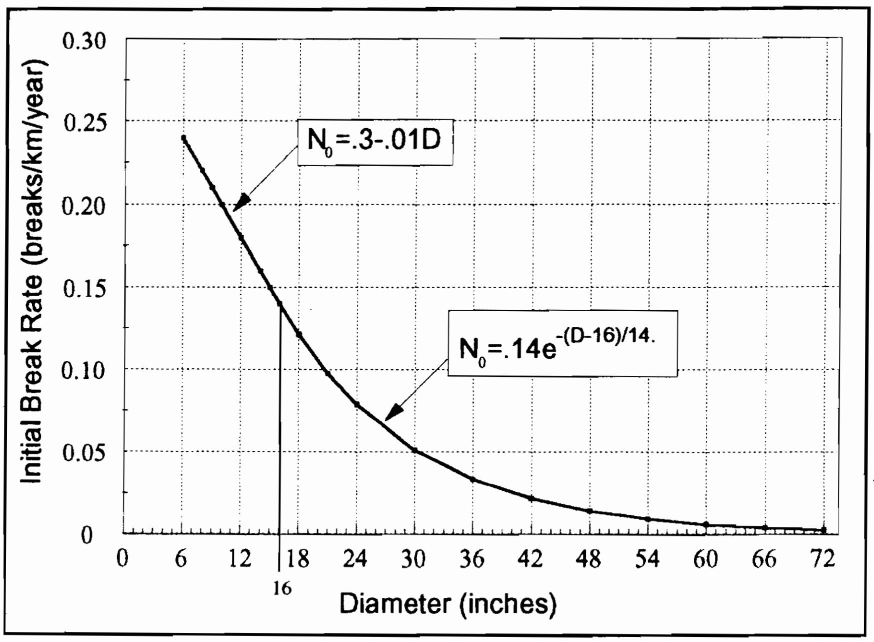

Figure 3.2 Decreasing Initial Break Rate Versus Diameter Relationship

The basic equation for break rate and growth of break rate in equation (3.6) are combined with the break rate versus diameter relationship of equation (3.7), plus the extension for larger diameter pipes to form the following modeled break rate as a function of time:

$$N(t,D) = N_0(D)\, e^{.1(t-t_0)} \equiv \begin{cases} (.3-.01D)\, e^{.1(t-t_0)} & \text{if } D \le 16 \\ .14\, e^{-(D-16)/14}\, e^{.1(t-t_0)} & \text{if } D \ge 16 \end{cases} \tag{3.8}$$

where

N = break rate (breaks/km/year)

N_0 = initial break rate

D = diameter of pipe (inches)

t = time (year)

t_0 = time of installation (year)

Notice that both the terms representing $N_0 = N_0(D)$ in the two cases of equation (3.8) have value 0.14 and slope -0.01 at D=16" and that N never reaches zero or takes on negative values. The break rates as modeled by equation (3.8) for a number of commercially available pipe diameters are plotted in Figure 3.3.

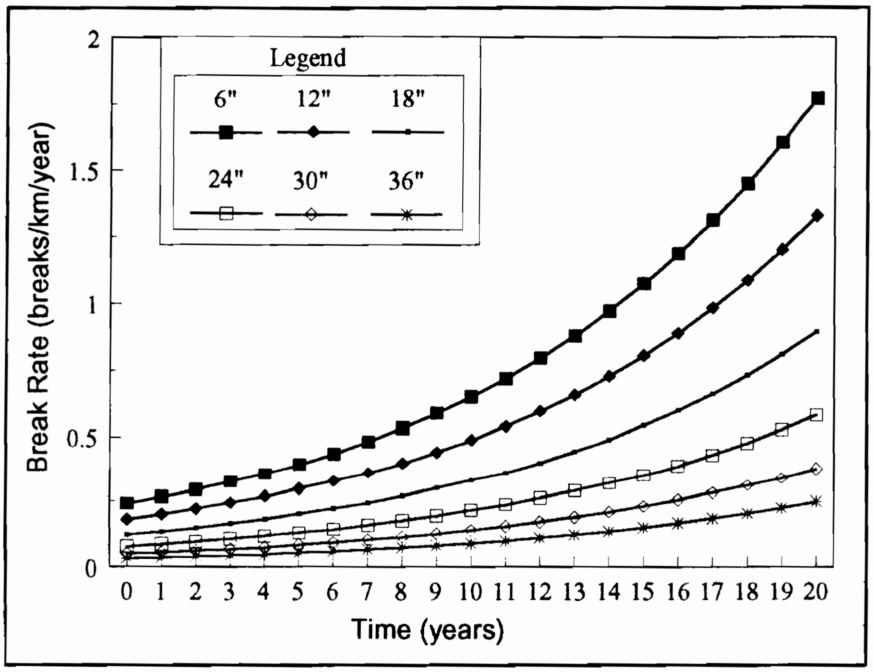

Figure 3.3 Break Rate as a Function of Time for Various Diameters

The above model will be used to project expected failure times for existing pipes and for their immediate or eventual replacement. Of course, real breaks will occur with great variability, and so the above model will only be used to approximate the annualized cost of future maintenance actions based on the predicted failure rates.

The expected time of the n^{th} future failure can be found by integrating equation (3.8) from t_{now} to t (multiplied by the length of the pipe segment), equating this to n (n=1 for expected time to the *next* failure), and solving for t.

Thus

$$\int_{t_{now}}^{t} L\, N_0(D)\, e^{.1(\tau - t_0)}\, d\tau = n \tag{3.9}$$

where

t = time (years) of n^{th} future failure (n=1 for next failure)

t_{now} = current time (year)

t_0 = time of installation (year)

L = length of the pipe segment (km)

$N_0(D)$ = initial break rate vs. diameter relationship as defined by eqn. (3.8)

n = number of future failures.

Solving equation (3.9) for t gives

$$t = \begin{cases} t_0 + 10 \ln\left(e^{.1(t_{now} - t_0)} + \dfrac{n}{10\, L\, (.3 - .01 D)}\right) & \text{if } D \le 16 \\ t_0 + 10 \ln\left(e^{.1(t_{now} - t_0)} + \dfrac{n\, e^{(D-16)/14}}{1.4\, L}\right) & \text{if } D \ge 16. \end{cases} \tag{3.10}$$

As an example of determining expected break times, consider a 16" diameter pipe segment of length .5 km. The expected time from installation to first failure is found by solving equation (3.10) with t_0=t_{now}=0, and n=1, yielding t=8.87 years. Now, given that the first failure occurred at 8.87 years, the expected time of failure for the second break is found by solving with t_0=0, t_{now}=8.87 and n=1 (or t_0=t_{now}=0 and n=2) to give t=13.50 years, with a 4.63 year expected time between the first and second breaks. Likewise, the expected time to failure for the third break is 16.65 years with a 3.15 year expected time between the second and third breaks.

We must also consider the case when the pipe for which we are estimating costs has been in place for several years and may have experienced previous breaks. We will assume that the break rate is still modeled by equation (3.8) regardless of the previous history of breaks. For example, suppose our 16" diameter pipe segment has had two previous breaks and is 12 years old (compare with the expected 13.5 years until the second break). We wish to compute the expected time of the next (third) break. We solve equation (3.10) with t_0=0, t_{now}=12 and n=1 for the time of first future (next) failure, giving t=15.58 years. Notice that the expected time of the third break is earlier than before, since the second break occurred earlier than expected. Also, the time between the second and third breaks of 3.58 years is longer than the 3.15 years before, since the failure rate is lower during the earlier years.

Before we can determine the expected cost of each repair, we must determine the capital replacement cost of installing a new pipe, since some maintenance actions will require replacing a physical section of pipe. The following relationship between the cost of replacing a pipe and its diameter was proposed by Quindry et al. (1981):

$$C_{CR} = cD^{r} \tag{3.11a}$$

where

C_{CR} = capital replacement cost (dollars/meter)

c = capital cost coefficient

D = pipe diameter (inches)

r = regression coefficient.

We will use Walski's (1984) refinement of the above model to include the dependence on the pipe construction material and range of diameters and typical coefficients for 1984 as follows:

where

C_{CR} = capital replacement cost (dollars/meter)

D = pipe diameter (inches).

$$C_{CR}(D) = \begin{cases} 14.1e^{.170D}, & \text{if } D \leq 8 \\ 3.00\,D^{1.40} & \text{if } 8 \leq D \leq 24 \\ 6.45\,D^{1.16} & \text{if } 24 \leq D \leq 48 \\ .656\,D^{1.75} & \text{if } D \geq 48 \end{cases} \tag{3.11b}$$

We have slightly modified these equations to include four diameter ranges, versus the original three, in order to allow for a more accurate representation of Walski's actual cost data as represented in his plot. The cost for small pipes (D≤8) are based on PVC pipes, the cost for medium pipes (8≤D≤48) are based on ductile iron pipes, and the cost for large pipes (D≥48) are based on concrete pipes. The diameters that we will consider as commercially available are 4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 21, 24, 30, 36, 42, 48, 54, 60, 66, and 72 inches. Their respective capital replacement costs per meter are plotted in Figure 3.4.

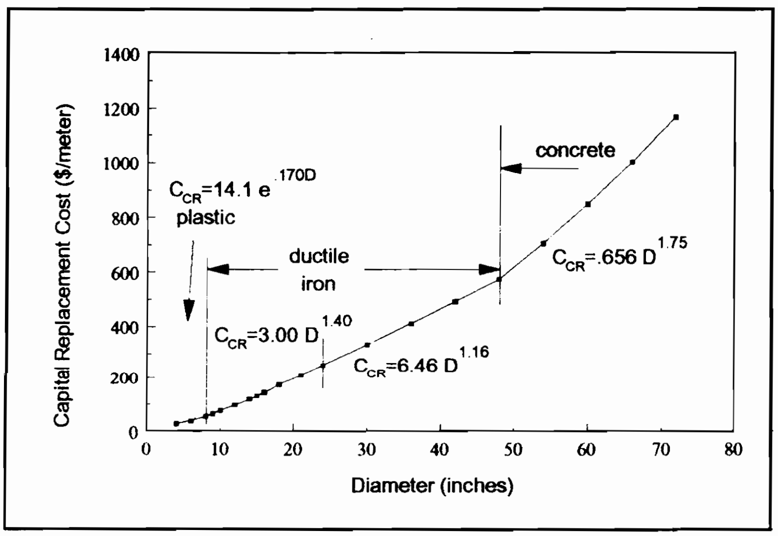

Figure 3.4 Capital Replacement Cost for Commercially Available Diameters

Now that a pipe failure model is in place and we have a model for the capital replacement cost for a new pipe, we can formulate a model for repair costs. Small leaks that are caused by a hole or small crack can be fixed with a repair clamp that wraps around the pipe, or for larger diameters, by welding a patch onto the pipe. One model reported by Walski (1984) that was useful for approximating the maintenance costs for repairing such a break in the Buffalo District from a U.S. Army Corps of Engineers study (1981) is as follows:

$$C_M = 600\ D^{0.40}. \tag{3.12a}$$

where

C_M = maintenance cost (dollars/break)

D = pipe diameter (inches),

and includes allowances for crew cost, equipment, sleeve, paving and tools.

Occasionally, larger longitudinal cracks or crushed pipes require the replacement of a physical section of pipe. We will assume this occurs 10% of the time, that sections are 10 meters long (these are variable parameters in the model), and that the costs incurred in equation (3.12a) still hold. Thus the average or expected maintenance repair cost for a break becomes

$$C_M = 600\ D^{0.40} + f_{SEC}\ L_{SEC}\ C_{CR}(D) \tag{3.12b}$$

where

f_{SEC} = fraction of cracks requiring section replacement (0.1)

L_{SEC} = length of physical section (10 meters)

$C_{CR}(D)$ is computed from equation (3.11b).

We use this model, but it is just one viable alternative. Each utility must determine the best model for local use and continue to upgrade it based on the most current information, adjusting it for current year dollars. Using the above formula, the estimated noninflated cost of repairing a break in our example of a 16" pipe is $600(16)^{.4}+.1(10)(3\cdot16^{1.4})$ = \$1964.

In order to compute and compare annualized prices, we use inflation-free real prices and real interest rates when discounting and annualizing costs. A formula for relating the market interest rate, real interest rate, and inflation rate presented by Shamir and Howard (1979) is

$$(1+r) = (1+R)(1+I) \tag{3.13}$$

where

r = market interest rate

R = real interest rate

I = inflation rate.

Assuming a 4 percent inflation rate, and an 8 percent market interest rate, the real interest rate computes to R = (1.08 /1.04) -1 = 3.85 percent. Since this real rate of interest already takes inflation into account, we do not need to adjust costs for inflation in our analysis.

Each cost outlay is annualized by first discounting the expense to the current year, and then converting it into an annualized cost using the following formulas from Grant, et al. (1987):

$$C_{PV} = \frac{C_{AC}}{(1+i)^y} \tag{3.14}$$

where

C_{PV} = cost discounted to present value (dollars)

C_{AC} = actual (uninflated) cost in year of the expense (dollars)

i = real interest rate (.0385)

y = year of expense,

and

$$C_{ANN} = C_{PV} \frac{i(1+i)^y}{(1+i)^y - 1} \tag{3.15}$$

where

C_{ANN} = annualized cost (dollars)

C_{PV} = cost discounted to present value (dollars)

i = real interest rate (.0385)

y = year of expense.

We can now estimate an optimal lifetime for each diameter of new pipe, based on the modeled capital replacement (installation) costs, failure rates, and maintenance repair costs. For each diameter of pipe, various lifetimes are considered at the expected failure times, and each candidate lifetime is analyzed by annualizing all costs incurred over that lifetime. As the lifetime is increased, the annualized costs decrease as is conceptualized in Figure 2.1 until maintenance costs begin to take over, and the annualized costs start to increase. We take this time to be the optimal life of the pipe based on financial considerations. (The options to overrule this decision time for replacement are discussed below). The optimal life and corresponding annualized cost for a 1000 m length of each of the twenty diameters considered to be commercially available are listed in Table 3.1.

Table 3.1 Optimal Life Characteristics for a New 1000 m Pipe

Diameter (inches)	Annualized Cost ($/meter)	Optimal Life (years)	Expected Breaks (per km)
4	2.77	22.9	23
6	3.59	24.7	26
8	4.61	26.8	30
9	5.22	27.9	32
10	5.82	28.9	34
12	7.03	31.2	39
14	8.23	33.1	42
15	8.83	34.1	44
16	9.42	35.2	46
18	10.60	37.2	49
21	12.40	40.3	54
24	14.24	43.1	58
30	17.21	48.1	63
36	20.12	53.0	67
42	23.04	57.9	71
48	25.99	62.7	75
54	30.92	68.0	83
60	36.30	73.2	91
66	42.07	78.2	98
72	48.25	83.3	106

These calculation are slightly dependent on length since we take into account the occasional section replacements required for longitudinal cracks and crushed pipes. Since the dependence on length is slight, we will assume that the cost for a 1000m length to be sufficiently representative for general use. Alternatively, one could adapt these computations to some representative length for each link in the network, say $L_{ij}/2$.

There are (at least) four important reasons for considering the replacement of a pipe segment. The first is to increase the hydraulic capacity of the pipe based on an expanded network and/or increased demand pattern. Second, a segment may be considered for replacement when the anticipated annualized costs of continuing to maintain it exceed the capital and future maintenance costs for a newly replaced pipe. Third, a utility may choose to replace pipe segments when the failure rate reaches a certain threshold level, regardless of economic consequences. This may be due to a limitation in the number of maintenance crews or due to a perceived inconvenience cost. This rate is a variable input parameter which we have arbitrarily chosen to be 10 breaks/year/km. Fourth, a segment may be removed based on the number of breaks in its life, regardless of time. This may simply be a conservative management practice that some utilities might use. This number is also a variable input parameter which we have chosen to be 20 breaks/km.

The annualized costs computed in Table 3.1 above are the cost coefficients that are input into the pipe network design submodel for new pipes (the c_{ijk}'s). We plot these in Figure 3.5 below.

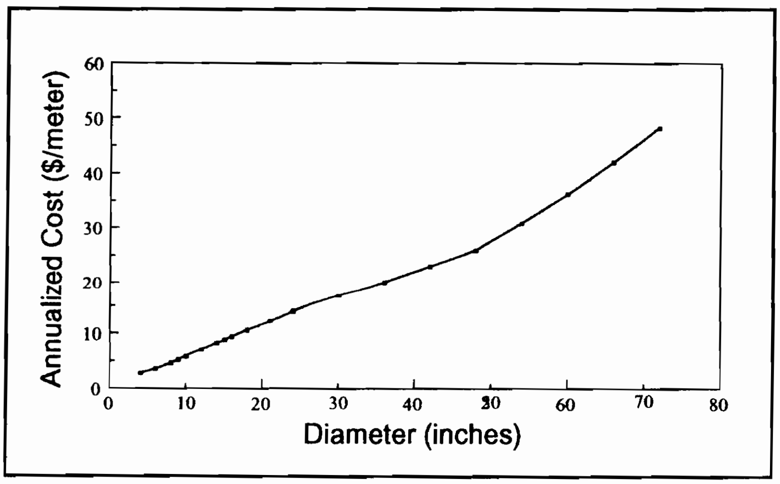

Figure 3.5 Annualized Cost versus Diameter

Perhaps the most important input into the decision of replace versus continue-to-maintain for an existing segment of fixed diameter pipe is "Which diameter for the new pipe should be considered?" Since the network is being expanded, and demand patterns may have increased, there is no reason to believe that a pipe of the same diameter will most cost effectively meet the new requirements, particularly since the friction loss coefficient will be different for the existing and new pipes. In the NOP formulation of Section 3.2.1, the X_{ijk} variables are used to indicate the length of existing pipe segment that is retained, with the x_{ijk} variables indicating the length of replacement segments and their corresponding diameters. However, the problem is more difficult to solve with this formulation and it does not make practical sense to remove part of a segment and leave the rest. If part of an old, failing, tuberculated pipe segment is going to be replaced, then all of it will probably be replaced in practice. Therefore, we will simplify this problem by deciding ahead of time in the pipe reliability submodel whether or not each segment will be replaced or not, essentially making the X_{ijk}'s fixed constants (0 or a_{ijk}) in the network optimization routine.

Now, in order to determine the diameter of pipe to be considered for replacement, we investigate the hydraulic properties of the existing pipe segment under its current flow pattern and compare the costs of retaining this segment with those of a comparable new pipe under a new baseline flow pattern. The new flow pattern to be considered will be the flow in the link for a representative demand pattern as if the whole expanded network were designed with all existing pipes in place, and with a new parallel pipe added along the entire length of each existing pipe. (This is step II in the integrated design approach from Section 3.1.) In this way, a representative or natural baseline flow for each link in the network under the new conditions is determined that takes into account the new demand along with the old network structure. (Alternatively, the submodel can be run from "scratch", that is, with all existing pipes removed, and new pipes in their place.) If, as a result, the link under consideration requires a pipe having a significantly greater hydraulic capacity than the existing pipe, then the pipe may need to be replaced regardless of the financial tradeoffs between replacement versus the continuation to maintain the pipe as discussed above. This replacement decision will not currently be quantified and will be considered as a subjective option to the designer. We will consider replacing a segment if a replacement pipe with an equal or lower hydraulic gradient (frictional head loss per unit length) is less expensive in annualized cost. In order to compute the hydraulic gradient, however, we need to estimate the Hazen-Williams coefficients for the existing and new pipes.

Walski (1984) determined that the roughness of a pipe (inversely proportional to C_{HW}) increases approximately linearly with time, with a change in slope at around thirty years. These changes in roughness depend on many factors, including pipe material, diameter, flow rate, and water composition, and, ideally could be measured pipe by pipe. An example chart of this relationship for a hypothetical 12" pipe under slight, appreciable, or severe corrosion conditions is given. Since the relationship with diameter is unknown, we simplify the model by using the appreciable corrosion curve and generalize this for all pipe diameters as follows:

$$C_{HW} = \begin{cases} 130-1.67t & \text{for } 0 \le t \le 30 \text{ years} \\ 80-.286(t-30) & \text{for } 30 \le t \le 100 \text{ years.} \end{cases} \tag{3.16}$$

This enables us to model the changing nature of C_{HW} for new and existing pipes, without adding unnecessary complexity. This relationship will be used to "age" all pipes under continue-to-maintain strategies 20 years before any hydraulic analysis is performed in order to ensure that the WDS will continue to operate as designed at least halfway through the 40 year financial time horizon. For example, our illustrative 16" pipe that is twelve years old is modeled in the current year with C_{HW}= 130-1.67(12) = 110, but for hydraulic analysis under a continue-to-maintain plan, this pipe will be aged to 32 years, with C_{HW}= 80-.286(32-30) = 79. Likewise, a new replacement pipe will not be analyzed with its initial C_{HW}=130, but with an aged C_{HW}= 130-1.67(20) = 97.

Now, we are prepared to determine the diameter of pipe to be considered for replacement. We choose the smallest pipe in the list of commercially available diameters that has less hydraulic gradient with the baseline flow than the existing pipe would have with the same flow. Suppose that for our example, we determine that the new baseline flow for the link is 600 m^3/hour. Recall that the Hazen-Williams coefficient for the existing and new pipes, respectively, are 79 and 97. The hydraulic gradient due to friction in the existing 16" pipe with a flow of 600 m^3/hour and C_{HW}=79 as determined by the Hazen-Williams equation (3.3c) is

$$\begin{aligned} \phi/length &= 7.6x10^3\, Q^2\, (C_{HW})^{-1.85} D^{-4.87} \\ &= 7.6x10^3\, (600.)^2\, (79.)^{-1.85} (2.54 \cdot 16)^{-4.87} \\ &= .0123 \ m/m \ (meters\ head\ loss/meter\ of\ pipe). \end{aligned}$$

The smallest diameter considered commercially available that has a lower hydraulic gradient than .0123 with the baseline flow of 600 m³/hour and C_{HW}=97 is 15". For this diameter, the hydraulic gradient is given as follows:

$$\begin{aligned}\phi / length &= 7.6x10^3\, Q^2\, (C_{HW})^{-1.85}\, D^{-4.87} \\ &= 7.6x10^3\, (600.)^2\, (97.)^{-1.85}\, (2.54 \cdot 15)^{-4.87} \\ &= .0115\ m/m\ (meters\ head\ loss/meter\ of\ pipe).\end{aligned}$$

A smaller new pipe works since it has less roughness. However, the optimization process may still choose a larger pipe if this submodel determines from the cost analysis below that the pipe will be replaced.

We can now compare the annualized construction and maintenance costs between replacing a pipe in the current year and with replacing it in some future year (at an expected break time). Continuing our example, we will compare the annualized cost over forty years of replacing our example 16" pipe segment that is 0.5 km long, and 12 years old in the current year, with a new 15" pipe for various lifetimes of the existing pipe. We will detail the analysis for replacement in the twentieth year (15th break), and summarize the results for all cases.

The expected maintenance times for the existing pipe are found by solving equation (3.10) with t_0=-12, t_{now}=0, and n=1,2,3,4 yielding 3.58, 6.21, 8.29, 10.0, ... years, each costing $1964 as we computed before. The annualized costs from Table 3.1 for the new 15" pipe are $8.83/m/yr. If we add up the discounted expected maintenance costs from the existing pipe (including the break in year 20) and discount the annualized costs of $4415/yr for the new pipe (from year 21 until the 40 year time horizon), we find that the present value of this option to replace during year 20 at the 15th break is $45,270, as illustrated below in Figure 3.6.

A cost analysis summary of the annualized costs for each time of expected failure is plotted in Figure 3.7. The optimal least cost year of replacement is 23 years, with a present value cost of $44,696 and annualized cost of $2208. The curve zigzags since we have a discrete cutoff of a forty year time horizon. For this option that replaces the pipe in year 23, there are 22 expected failures and corresponding maintenance actions for the existing 16" pipe before its replacement in year 23, followed by 2 expected failures and

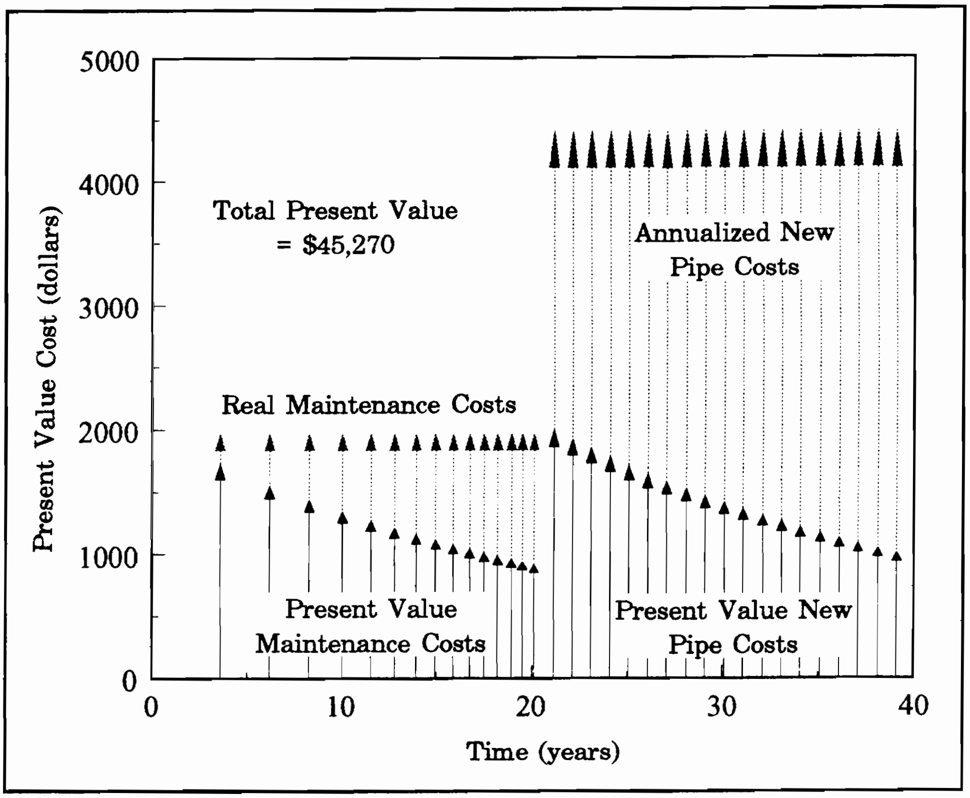

Figure 3.6 Present Value Costs for Replacement in Year 20

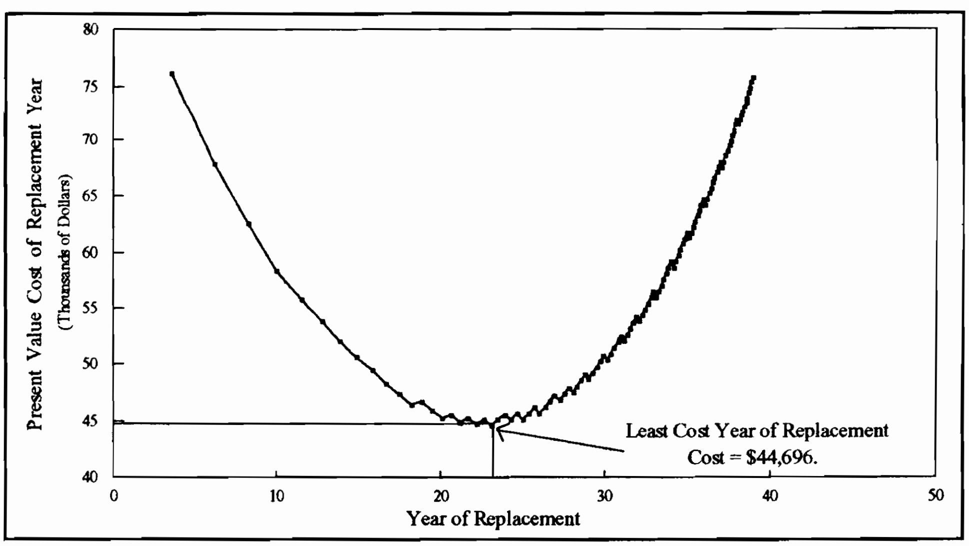

Figure 3.7 Determining the Least Cost Year of Replacement

corresponding maintenance actions for the new 15" pipe between year 23 and year 40. (However, we just use the annualized cost from Table 3.1 for the analysis.) Since the optimal replacement time does not occur during the next five years, we conclude that this existing pipe segment will *not* be a candidate for replacement in the current phase of network upgrading unless the network design is later determined to be hydraulically infeasible due to this segment. This assumes an approximate five year budgetary cycle for network upgrades. If the optimal year had been during this five year period, then this segment would have been scheduled for replacement during the current cycle, and prioritized based on a cost savings ranking among all segments.

This concludes our example to demonstrate the steps involved in the pipe reliability and cost analysis submodel. These steps are summarized below.

<u>Summary of Pipe Reliability and Cost Analysis Submodel Steps</u>. For each existing pipe segment of fixed diameter and Hazen-Williams coefficient, do the following:

(1) Compute the expected future break times during the next forty years for the existing pipe using equation (3.10), and the corresponding maintenance cost (uninflated) for each failure of the existing pipe using equation (3.12b).

(2) Compute the Hazen-Williams coefficient for the existing pipe after twenty years of *additional* aging using equation (3.16).

(3) Using the baseline maximum flow rate resulting from the previous optimization submodel, choose a candidate replacement pipe diameter based on the smallest diameter in the list of those commercially available that has less hydraulic gradient (computed from equation 3.3) than the existing pipe would have using the same flow rate. For each candidate diameter, use a Hazen-Williams coefficient after twenty years initial aging using equation (3.16).

(4) Compute the expected capital replacement cost (uninflated) for the new replacement pipe using equation (3.11b).

(5) Compute the expected future break times during the remainder of the forty year horizon for the new pipe using equation (3.10), and the corresponding maintenance cost (uninflated) for each failure of the new pipe using equation (3.12b).

(6) For each candidate replacement year from 0 to 40 years, compute the present value and annualized costs of the replacement and maintenance schedules determined above in steps (1),(5) and (6) using equations (3.14) and (3.15), and find the minimum annualized (or present value) cost to determine the optimal year of replacement, as was done in Figure 3.7.

3.4 Model Parameters

Several key quantities introduced in this chapter are variable input parameters for their respective submodels. These are summarized below.

1. Required percent of hydraulic pressure that should be maintained for each demand pattern flow in the redundancy analysis (80%).

2. The water flow rate Q_{iik}' at which C_{HW} was measured for each existing pipe segment (100 meters^3/hour).

3. Percent of repairs requiring replacement of a physical section of pipe (10%).

4. Length of physical pipe sections for each diameter (10 meters).

5. Critical break rate at which pipe segments are automatically scheduled for replacement (10 breaks/year/km).

6. Critical number of breaks at which pipe segments are automatically scheduled for replacement (20 breaks/km).

4. DECOMPOSITION ALGORITHM FOR THE NETWORK OPTIMIZATION PROBLEM

In this chapter we develop the theoretical basis for a decomposition scheme to solve the network optimization problem NOP and then discuss various related algorithmic and implementation issues. The manipulation of problem NOP described below is similar in spirit to that used by Collins et al. (1978), in that it reveals a critical convex cost network flow programming subproblem but is more complicated because of the presence of the head bounds (3.5g), and because we do not assume that the direction of flow in each pipe is known *a priori*. The motivation behind this manipulation is to decompose the problem into a master problem that essentially represents a projection of the original problem onto the space of the design variables $w \in W$, and a suitable subproblem. This subproblem turns out to be a specially structured convex network flow programming problem that can be easily solved. The idea, then, is to solve the original problem through the solution of a sequence of subproblems, as directed by an iteration update of the design variables via the master problem. In this process, we also develop a novel variable-dimension reduction strategy that greatly facilitates the handling of large-scale instances of the problem.

4.1 Structural Manipulation and Reformulation of the Network Optimization Problem

In order to develop the proposed decomposition scheme, we need to first restructure the problem into an equivalent, convenient format. Toward this end, let us construct two additional transshipment nodes $p \equiv (n+1)$ and $q \equiv (n+2)$ and create new arcs $(p,0)$, $(0,q)$, (i,p) for each $i \in D$, and (q,i) for each $i \in D$. Designate the flows on these arcs as y_{p0}, y_{0q}, y_{ip} for $i \in D$, and y_{qi} for $i \in D$, respectively. Figure 4.1 depicts a generic network flow diagram with the source and demand nodes schematically grouped together for the purpose of illustration. The motivation for this construction will become evident when we establish the theoretical properties of the resultant modified problem. As we shall show, for every fixed set of design variables $w \equiv (x, X, H_s) \in W$, the residual problem turns out to be a convex cost network flow programming problem that yields unique values for the flows in the network depicted in Figure 4.1. In fact, for this network flow problem, given any set of partial flows $Q_{ij} > 0$, $(i,j) \in A$, that satisfy just the supply constraints for the nodes $i \in S$, we shall see that the other flows Q_{i0} for $i \in S$, y_{ip} and y_{qi} for $i \in D$, y_{p0}, and y_{0q} are all uniquely

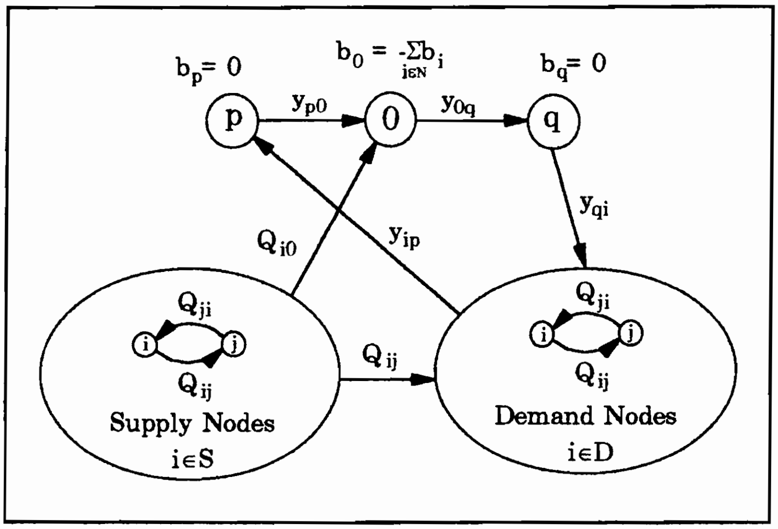

Figure 4.1 Generic Network Flow Diagram

determined to yield a feasible completion to the flow vector. Hence, this construction permits us to view the aforementioned network problem in the space of the Q_{ij}, $(i,j) \in A$ variables and over a relaxed region characterized by the supply constraints alone. Now, given the established network links, let us rewrite the flow conservation and logical constraints (3.5b) - (3.5d) and (3.5j) - (3.5k) become the following

$$\sum_{j:(i,j) \in A} Q_{ij} - \sum_{j:(j,i) \in A} Q_{ji} + Q_{i0} = b_i \quad \text{for each } i \in S \tag{4.1a}$$

$$\sum_{j:(i,j) \in A} Q_{ij} - \sum_{j:(j,i) \in A} Q_{ji} + y_{ip} - y_{qi} = b_i \quad \text{for each } i \in D \tag{4.1b}$$

$$y_{0q} - y_{p0} - \sum_{i \in S} Q_{i0} = b_0 \quad \text{for node } 0 \tag{4.1c}$$

$$y_{p0} - \sum_{i \in D} y_{ip} = 0 \text{ for node } p \tag{4.1d}$$

$$\sum_{i \in D} y_{qi} - y_{0q} = 0 \text{ for node } q \tag{4.1e}$$

$$Q_{ij} \geq 0 \text{ for } (i,j)\in A,\ Q_{i0} \geq 0 \text{ for } i\in S,\ y_{p0} \geq 0,\ y_{0q} \geq 0,$$
$$y_{ip} \geq 0 \text{ for each } i\in D,\ y_{qi} \geq 0 \text{ for } i\in D \tag{4.1f}$$

Define the revised overall flow vector $Q_R \equiv (Q,y)$ to have components $\{Q_{ij}$ for $(i,j)\in A$, Q_{i0} for $i\in S$, y_{p0}, y_{0q}, y_{ip} for $i\in D$, and y_{qi} for $i\in D\}$, and denote the reformulated feasible region as $Q_{FR} = \{Q_R$: constraints (4.1a) - (4.1f) are satisfied$\}$. Accordingly, consider the following subproblem defined for a fixed $w = (x,X,H_s)$ in W.

SP(w): Minimize

$$F(Q_R,w) = \sum_{(i,j)\in A} \int_0^{Q_{ij}} \phi_{ij}(q,x_{ij\cdot},X_{ij\cdot})dq + \sum_{i\in D}(H_{iU}y_{ip} - H_{iL}y_{qi}) + \sum_{i\in S}(F_i + H_{si})Q_{i0} \tag{4.2a}$$

subject to

$$Q_R \in Q_{FR} \tag{4.2b}$$

Observe that problem SP(w) is a convex cost network flow programming problem in the variables $(Q_{ij}, Q_{io}, y_{ip}, y_{qi}, y_{p0}, y_{0q})$ and can be solved very efficiently by exploiting the underlying network structure. Moreover, the following two results reveal the relevance of SP(w) in solving the pipe network design problem NOP.

Theorem 1. For any fixed $w\in W$, Problem SP(w) is feasible and bounded in value, and has a unique optimal solution. Hence, this optimal solution can be designated as a function $Q_R^*(w)=[Q^*(w),y^*(w)]$ of $w\in W$.

Proof: Let $w=(x,X,H_s) \in W$ be fixed. It is readily verified that SP(w) is feasible by considering the solution $Q_{ij}=0$ for all $(i,j)\in A$, $y_{ip}=0$ for all $i\in D$, $y_{p0}=0$, $Q_{i0}=b_i$ for each $i\in S$, $y_{qi}= -b_i$ for all $i\in D$, and $y_{0q} = -\sum_{i\in D} b_i$. Next, we show that SP(w) is bounded in value. Define

$$\Gamma = \{(Q_{ij}) \geq 0,\ (i,j)\in A: \sum_{j:(i,j)\in A} Q_{ij} - \sum_{j:(j,i)\in A} Q_{ji} \leq b_i \ \forall i\in S \tag{4.3}$$

and consider a projection of SP(w) onto Γ.

For any $(Q_{ij}) \in \Gamma$, define the residual flow $r_i \equiv r_i(Q) \equiv b_i - \sum_{j:(i,j)\in A} Q_{ij} + \sum_{j:(j,i)\in A} Q_{ij}$ for each $i \in N$, and note that $\sum_{i\in N} r_i = \sum_{i\in N} b_i = -b_0$. Now, given any $(Q_{ij}) \in G$, in order to satisfy Eqs. (4.1b) and (4.1f) we must find (nonnegative) y_{ip} and y_{qi} $\forall i \in D$ such that $y_{ip} - y_{qi} = r_i$ $\forall i \in D$. Having done this, Eq. (4.1a) will uniquely yield $Q_{i0} = r_i \geq 0$ $\forall i \in S$, Eq. (4.1d) will uniquely yield $y_{p0} = \sum_{i\in D} y_{ip} \geq 0$, and Eq. (4.1e) will uniquely yield $y_{0q} = \sum_{i\in D} y_{qi} \geq 0$. Moreover, Eq. (4.1c) will be automatically satisfied since $y_{0q} - y_{p0} - \sum_{i\in S} Q_{i0} = \sum_{i\in D}(y_{qi} - y_{ip}) - \sum_{i\in S} r_i = -\sum_{i\in N} b_i = b_0$ from above. Accordingly, SP(w) can be equivalently restated as the following problem

$$\begin{aligned}
\underset{(Q_{ij})\in\Gamma}{\text{Minimize}} \sum_{(i,j)\in A} \int_0^{Q_{ij}} \phi_{ij}(q, x_{ij\cdot}, X_{ij})\,dq + \sum_{i\in S}(F_i + H_{si}) r_i \\
+ \text{Minimum}\{\sum_{i\in D}(H_{iU} y_{ip} - H_{iL} y_{qi}) : \; y_{ip} - y_{qi} = r_i,\; y_{ip} \geq 0,\; y_{qi} \geq 0 \\
\text{for each } i \in D\}
\end{aligned} \tag{4.4}$$

where $r_i \equiv (b_i - \sum_{j:(i,j)\in A} Q_{ij} + \sum_{j:(j,i)\in A} Q_{ji})$ $\forall i \in N$. Substituting $y_{qi} = (y_{ip} - r_i)$ for each $i \in D$ from the equality constraint, the inner minimization problem in Eq. (4.4) further reduces to the following trivial separable problem

$$\begin{aligned}
&\underline{\text{Minimize}} && \sum_{i\in D}[(H_{iU} - H_{iL}) y_{ip} + H_{iL} r_i] \\
&\underline{\text{subject to}} && y_{ip} \geq r_i \text{ for } i \in D \quad (\text{from } y_{ip} - r_i = y_{qi} \geq 0) \\
& && y_{ip} \geq 0 \text{ for } i \in D
\end{aligned}$$

which has the <u>unique</u> solution $y_{ip} = \max\{0, r_i\}$ $\forall i \in D$, having objective function value $\sum_{i\in D:\, r_i \geq 0} H_{iU} r_i + \sum_{i\in D:\, r_i < 0} H_{iL} r_i = \sum_{i\in D}[H_{iL} r_i + (H_{iU} - H_{iL}) \cdot \max\{0, r_i\}]$. Substituting into Eq. (4.4) we get

$$\underline{\mathrm{SP(w)}}: \text{ Minimize}[\sum_{(Q_{ij})\in\Gamma} \sum_{(i,j)\in A} \int_0^{Q_{ij}} \phi_{ij}(q, x_{ij}, X_{ij})dq + \sum_{i\in S}(F_i + H_{si})r_i$$

$$\sum_{i\in D}[H_{iL}r_i + (H_{iU} - H_{iL})\cdot\max\{0, r_i\}] \tag{4.5}$$

$$\text{where } r_i \equiv (b_i - \sum_{j:(i,j)\in A} Q_{ij} + \sum_{j:(j,i)\in A} Q_{ji})\forall i\in N.$$

But note that $[\sum_{i\in S}(F_i + H_{si})r_i + \sum_{i\in D: r_i\geq 0} H_{iU}r_i + \sum_{i\in D: r_i<0} H_{iL}r_i] \geq [H_L\sum_{i\in N} r_i] = [H_L\sum_{i\in N} b_i] \geq 0$ (by Eq. (3.2)). Hence, the objective function of SP(w) in Eq. (4.5) is nonnegative, implying that SP(w) has a bounded minimal value.

Finally, let us consider the uniqueness of the optimal solution to SP(w). Note that the first term of Eq. (4.5) is given by

$$\sum_{(i,j)\in A} \int_0^{Q_{ij}} \phi_{ij}(Q_{ij}, x_{ij}, X_{ij}) = \sum_{(i,j)\in A} (\sum_{k=1}^{K} 8.21\times 10^3 (Q_{ij}/C_{HWx_{ijk}})^{2.852} d_k^{-4.87} x_{ijk}$$
$$+ \sum_{k=1}^{K} 8.21\times 10^3 (Q_{ij}/C_{HWX_{ijk}})^{2.852} d_k^{-4.87} X_{ijk}$$

which is of the form $\Sigma a_{ij}Q_{ij}^{2.852}$, where the a_{ij}'s are positive constants, and is hence strictly convex in all the (Q_{ij}) variables. Furthermore, the second term of Eq. (4.5) is linear, hence convex, in (Q_{ij}). Finally, the rest of the terms in Eq. (4.5) are convex since r_i is linear in (Q_{ij}) and $(H_{iU}-H_{iL})\cdot\max\{0,r_i\}$ is convex with $(H_{iU}-H_{iL})>0$ for all $i\in D$. Hence Eq. (4.5) is strictly convex. Thus Eq. (4.5) has a unique optimal solution $(Q_{ij})^*$. As shown above, this will necessarily yield unique accompanying values for Q_{i0}^* for $i\in S$, y_{ip}^* and y_{qi}^* for $i\in D$, y_{p0}^*, and y_{0q}^*. Thus SP(w) has a unique optimal solution which can be designated as $Q^*(w)$, $y^*(w)$. ■

<u>Theorem 2</u>. Consider a feasible solution $(\bar{w}, \bar{Q}, \bar{H})$ to problem NOP. Then $Q^*(\bar{w}) = \bar{Q}$ and $y^*(\bar{w}) = 0$ uniquely solves $SP(\bar{w})$. Moreover, $(\bar{H}_i + E_i)$, $i\in N$, are an optimal set of Lagrange multipliers associated with constraints Eqs. (3.5c) and (3.5d) for the problem $SP(\bar{w})$.

Conversely, let $\bar{w} \in W$ be such that $y^*(\bar{w}) = 0$ at optimality in SP($\bar{w}$). Denote by v_i, $i \in N$, the set of Lagrange multipliers associated with constraints Eqs. (3.5c) and (3.5d) as obtained at optimality in SP($\bar{w}$), and define $\bar{Q} \equiv Q^*(\bar{w})$ and $\bar{H}_i \equiv v_i - E_i$ for all $i \in N$. Then $(\bar{w}, \bar{Q}, \bar{H})$ is feasible to problem NOP.

Proof: The KKT conditions for SP(w) are the primal feasibility conditions Eqs. (4.1a) - (4.1f), plus the dual feasibility and complementary slackness conditions as specified below.

$$\phi_{ij}(Q_{ij}, x_{ij\cdot}, X_{ij\cdot}) - v_i + v_j \geq 0 \text{ with equality if } Q_{ij} > 0 \quad \forall (i,j) \in A \tag{4.6a}$$

$$(F_i + H_{Si}) - v_i \geq 0 \text{ with equality if } Q_{i0} < 0 \quad \forall i \in S \tag{4.6b}$$

$$H_{iu} - v_i + v_p \geq 0 \text{ with equality if } y_{ip} < 0 \quad \forall i \in D \tag{4.6c}$$

$$-H_{iL} + v_i - v_q \geq 0 \text{ with equality if } y_{qi} < 0 \quad \forall i \in D \tag{4.6d}$$

$$- v_p \geq 0 \text{ with equality if } y_{p0} < 0 \tag{4.6e}$$

$$+ v_q \geq 0 \text{ with equality if } y_{0q} < 0 \tag{4.6f}$$

Note that since all v's appear as differences in Eqs. (4.6a) - (4.6f) we can arbitrarily fix $v_0 \equiv 0$. Now, consider a feasible solution $(\bar{w}, \bar{Q}, \bar{H})$ to NOP. Let $Q = \bar{Q}$, $y = \bar{y} = 0$, $\bar{v}_i = -(\bar{H}_i + E_i)$ $\forall i \in N$, $\bar{v}_0 = 0$, $\bar{v}_q = 0$. Then noting that $\phi_{ij}(0, \bullet, \bullet) \equiv 0$ we have Eq. (3.5e) implying Eq. (4.6a), Eq. (3.5f) implying Eq. (4.6b), Eq. (3.5g) implying Eqs. (4.6c) and (4.6d), and furthermore, Eqs. (4.6e) - (4.6f) hold trivially. Moreover $(\bar{Q}, \bar{y})$ is primal feasible. Hence, since SP is a convex program, $(\bar{Q}, \bar{y})$ solves SP($\bar{w}$) with $\bar{v}_i$, $i \in N$ being optimal Lagrange multipliers associated with the constraints Eqs. (4.1a) and (4.1b). Also, by Theorem 1, $(\bar{Q}, \bar{y})$ is the unique optimum.

Conversely, let $\bar{w} \in W$, and let SP($\bar{w}$) yield $\bar{Q}$ and $\bar{y} = 0$ as the unique optimum, along with Lagrange multipliers $\bar{v}$ satisfying Eqs. (4.6a) - (4.6f) with $\bar{v}_0 \equiv 0$. By Eqs. (4.1a)

- (4.1c), we have Eqs. (3.5b) - (3.5d) holding at this solution $\bar{Q}$. Also, since $\bar{w} \in W$, we have Eqs. (3.5h), (3.5i), and (3.5k) holding true. Furthermore, since $\bar{Q}_{ij} \leq Q_{ij}^U$ by assumption, Eq. (3.5j) holds true. Hence, we need to verify that Eqs. (3.5e), (3.5f), and (3.5g) also hold true.

Constraint Eq. (3.5e) follows from Eq. (4.6a) since $(H_i+E_i)-(H_j+E_j) = \bar{v}_i - \bar{v}_j \leq \phi_{ij}(Q_{ij}, x_{ij}, X_{ij})$ with equality holding if $\bar{Q}_{ij} > 0$, and where $\phi_{ij}(Q_{ij}, x_{ij}, X_{ij}) = 0$ if $\bar{Q}_{ij} = 0$. Constraint Eq. (4.6b) implies constraint Eq. (3.5f) when $\bar{H}_i = \bar{v}_i - E_i$. Furthermore, by Eqs. (4.6c) and (4.6e), we get $\bar{H}_i + E_i = \bar{v}_i \leq H_{iU} + v_p \leq H_{iU}$, and by Eqs. (4.6d) and (4.6f), we get $\bar{H}_i + E_i = \bar{v}_i \geq H_{iL} + v_q \geq H_{iL}$, and so, Eq. (3.5g) holds. Hence, $(\bar{w}, \bar{Q}, \bar{H})$ is feasible to problem NOP, and this completes the proof. ■

Theorem 2 holds the key to projecting Problem NOP onto the space of the design variables $w \in W$. It states that any $\bar{w} \in W$ is (part of) a feasible solution in NOP if and only if $y^*(\bar{w}) = 0$ at optimality in SP($\bar{w}$). Noting this property, and the fact that y=0 in Eqs. (4.1a) - (4.1f) if and only if $y_{p0} = y_{oq} = 0$, consider the following projection PNOP of problem NOP onto the design variables w, where $\mu > 0$ is some penalty parameter:

$$\text{PNOP:} \quad \underline{\text{Minimize}} \; f_\mu(w) = [cx + CX + c_s H_s] + \mu [y_{p0}^*(w) + y_{0q}^*(w)] \tag{4.7a}$$

$$\underline{\text{subject to}} \quad w \equiv (x, X, H_s) \in W \tag{4.7b}$$

The alternate use of a quadratic penalty function for (4.7a) is also discussed in Section 4.3. The principal function of Problem PNOP is embodied in the following theorem, which is a consequence of the foregoing results for SP(•).

Theorem 3. Suppose that for some penalty parameter $\mu > 0$, we obtain an optimal solution w^* to Problem PNOP such that $y_{p0}^*(w^*) = y_{0q}^*(w^*) = 0$. Denote $Q^* = Q^*(w^*)$, and let H^*+E be a set of Lagrange multipliers associated with Eqs. (4.1a) and (4.1b) at optimality in SP(w^*). Then (w^*, Q^*, H^*) solves Problem NOP.

Proof: From Eqs. (4.1d) and (4.1e) of SP(w^*), we find that $y^*_{ip}(w^*) = y^*_{qi}(w^*) = 0$, for all $i \in D$ since $y^*_{p0}(w^*) = y^*_{0q}(w^*) = 0$. Hence, $y^*(w^*)=0$, and so, by the converse statement of Theorem 2, (w^*,Q^*,H^*) is feasible to NOP. Consider any other feasible solution $(\bar{w}, \bar{Q}, \bar{H})$ to NOP. By the first statement of Theorem 2, $y^*(\bar{w}) = 0$, which implies that $\bar{w}$ is feasible to PNOP having an objective value $c\bar{x} + C\bar{X} + c_s\bar{H}_s$. Since w^* is optimal to PNOP, we have, $cx^* + CX^* + c^s H^*_s \leq c\bar{x} + C\bar{X} + c_s\bar{H}_s$. Therefore, ($w^*,Q^*,H^*$) solves NOP, and this completes the proof. ■

Observe that if there exists some network configuration that satisfies the reliability requirements, then Problem PNOP is always feasible, irrespective of whether Problem NOP is feasible or not. Moreover, PNOP is a convenient decomposition of Problem NOP in which the special structure of the latter problem has been suitably exploited and a parameter μ has been used to penalize the total violation in demand corresponding to $Q_R^*(w)$. The problem has been effectively projected onto the space of the design variables $w \in W$ and the determination of flows and pressure heads for a fixed $w \in W$ has been reduced to the solution of a convex cost network flow programming problem SP(w) that can be solved very efficiently. In particular, note that this latter capability replaces the Hardy-Cross solver method, or the Newton-Raphson method, or the linear theory method, or other suitable numerical techniques used to determine (Q,H) in Problem NOP for a fixed $w \in W$ (see Wood and Charles, 1972). More importantly, by examining the variational properties of the specially structured problem SP($\cdot$), we can extract useful information on how to modify a given set of design parameters in order to improve the objective function value. Hence, our approach provides a novel and insightful alternative to the methods variously discussed in Jacoby (1968), Alperovits and Shamir (1977), Quindry et al. (1979), Gessler (1985), Lansey and Mays (1985), Ormsbee (1985), and Fujiwara et al. (1987) for analyzing networks with given layouts, and in Rowell and Barnes (1982), Morgan and Goulter (1985), Su et al. (1987), and Loganathan et al. (1990) for analyzing distribution systems that include the problem of selecting the network layout.

4.2 A Dimension Reduction Via the Definition of Composite Variables for PNOP

The objective of the network optimization problem NOP is to find the least cost permissible design solution $w^*=(x,X,H_a)^* \epsilon W$ that admits a set of flows which satisfy demands along with associated pressure head requirements. The projection PNOP of this problem onto the design space of the w-variables permits us to adjust $w \epsilon W$ such that the resultant completion of the overall solution to NOP, as determined by the subproblem SP, yields the best objective value.

In general, there will be a large number of x_{ijk} and X_{ijk} variables since there can be many pipe links (i,j), and each one can be composed of multiple existing and new segments using a number of candidate pipe diameters d_k=1,...,K. Thus the dimensionality of the problem PNOP can become prohibitively large if several such candidates are to be considered.

In order to simplify the search process, we fix the X variables in the reliability and cost model as outlined in Section 3.3 so that constraint (3.5i) becomes

$$\sum_{k=1}^{K} x_{ijk} = L_{ij} - \sum_{k=1}^{K} X_{ijk} \text{ for each } (i,j)\epsilon A,\ i<j, \tag{4.8}$$

where the right-hand side terms are now all constants. We can now define an effective length L'_{ij} for each link to be the remaining length of the link to be designed using newly constructed pipe segments as follows:

$$L'_{ij} = L_{ij} - \sum_{k=1}^{K} X_{ijk} \text{ for each } (i,j)\epsilon A,\ i<j. \tag{4.9}$$

One traditional method for coping with the size of the problem by reducing the number of x_{ijk} variables is to limit the number of candidate diameters from which to select, perhaps choosing different candidates for each pipe based on some prior analysis of the system (see, for example, Alperovits and Shamir, 1977). This method is motivated by the fact that, under certain reasonable assumptions on pipe costs (see Rowell, 1979), each link will always have either a single diameter segment or some two segments having adjacent diameters from the set of standard available diameters. These assumptions require strict convexity in the cost coefficients as a function of the diameters. Hence, using this property, the candidate list for each pipe is heuristically restricted to span a narrow range of consecutive, admissible diameters in order to reduce the search space.

For example, suppose that an optimal link design requires segment diameters of 16" and 18", say, each of equal length. If fortuitously, we have considered a candidate list of diameters as (14", 16", 18", 21"), we will recover the two foregoing diameters at optimality, but will have reduced the number of variables considered for this link from 20 possibilities to only 4. However, we do not know beforehand which diameters in the overall candidate list will be selected as being optimal for each link. So, in our example, if the list of restricted candidates was selected as (18", 21", 24", 30"), then the overall optimal choice would not be achievable, and the algorithm would likely select an 18" diameter pipe for the whole length, a suboptimal result. Hence, unless a large range of candidate diameters is selected for each link, the optimization process will have to be repeated using different candidate lists for each link, perhaps many times, before the solution obtained selects diameters for each pipe segment that do not lie at the extreme ends of its permissible range.

As an alternative, we recommend a new approach that significantly reduces the dimensionality of the problem imposed by several candidate diameters, while not restricting the consideration of such candidates. This is achieved by defining a set of composite variables χ_{ij}, $(i,j)\in A$, that represent the x_{ijk} variables $\forall i,j,k$ in a compact form. The property of the solution that we exploit in this transformation is that at most two distinct diameter segments need to be considered for composing each pipe and that these diameters only need to be adjacent to each other. Hence, instead of selecting from a group of x_{ijk} variables for each segment (i,j) in order to satisfy constraints (3.5i) of PNOP, we let χ_{ij} represent in a "sliding" fashion the joint decision on which adjacent pair of available pipe diameters is used, and how much of the length (L'_{ij}) is assigned to each of these diameter segments.

Specifically, let us introduce the dummy variables λ_{ijk} $\forall$ i,j,k to denote the fraction of pipe length L'_{ij} that is comprised by x_{ijk}. Hence, for each $(i,j)\in A$,

$$\sum_{k=1}^{K} \lambda_{ijk} = 1,\ \lambda_{ijk} \geq 0\ \forall k=1,\ldots,K, \text{ and } x_{ijk} \equiv L'_{ij}\, \lambda_{ijk}\ \forall k=1,\ldots,K. \tag{4.10}$$

Now, let us define the composite variables

$$\chi_{ij} = \sum_{k=1}^{K} k L'_{ij} \lambda_{ijk} \equiv \sum_{k=1}^{K} k x_{ijk}, \quad \text{where} L'_{ij} \leq \chi_{ij} \leq K L'_{ij}. \tag{4.11}$$

Given any x_{ijk} values, χ_{ij} is uniquely determined via Eq. (4.11). More importantly, under the aforementioned adjacency property satisfied by the λ-variables, the converse is also true, that is, given any χ_{ij} value in $[L'_{ij}, KL'_{ij}]$, given that at most two variables λ_{ijk}, $k \in \{1,\ldots,K\}$ can be positive and that these correspond to adjacent k values, Eq. (4.11) yields λ_{ijk} values uniquely and Eq. (4.10) uniquely gives the corresponding x_{ijk}-values. Specifically, the following solution is obtained (see Figure 4.2):

$$x_{ijk} = \begin{cases} 0 & \text{for } L'_{ij} \leq \chi_{ij} \leq (k-1)L'_{ij} \\ \chi_{ij} - (k-1)L'_{ij} & \text{for } (k-1)L'_{ij} \leq \chi_{ij} \leq kL'_{ij} \\ (k-1)L_{ij} - \chi'_{ij} & \text{for } kL'_{ij} \leq \chi_{ij} \leq (k+1)L'_{ij} \\ 0 & \text{for } (k+1)L'_{ij} \leq \chi_{ij} \leq KL'_{ij} \end{cases} \quad \forall\ (i,j) \in A,\ k=1,\ldots,K. \tag{4.12}$$

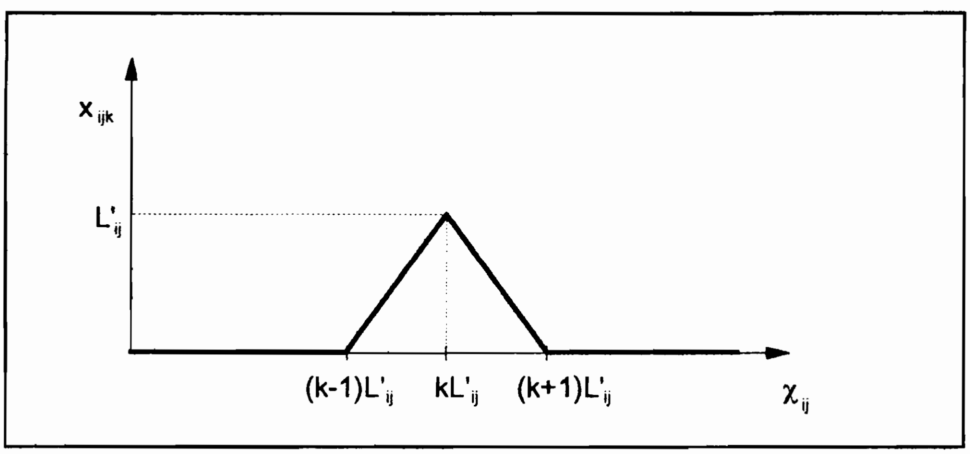

Figure 4.2 Inverse transformation of the composite χ_{ij} variables

Once these χ-variables are introduced, the full range of available diameters can be made candidates for each link (i,j) with a single χ_{ij} variable without any loss of representation. For example, if the diameters from Chapter 3 (4", 6", 8", 9", 10", 12", 14", 15", 16", 18", 21", 24", 30", 36", 42", 48", 54", 60", 66", 72") are available for a link (i,j) having L'_{ij}=1000m, then χ_{ij} can "slide" anywhere between 1000 (all 4" pipe) and 20000 (all 72" pipe). As χ_{ij} increases from 1000, less of the 4" pipe is selected and more of the 6" pipe is selected. For instance, χ_{ij}=1200 selects 800m of the 4" pipe and 200 m of the 6" pipe. When χ_{ij} increases up to 2000, all of the 1000m length is selected as a 6" pipe. In this fashion, twenty continuous x_{ijk} variables are reduced to one continuous χ_{ij} variable and the variables remain simply upper and lower bounded.

Note, however, that the inverse transformation (4.12) (see Figure 4.2), is neither linear nor differentiable. This means that the objective functions of problems NOP and PNOP are nondifferentiable functions of the χ_{ij} variables. In fact, with our assumption on pipe costs, the linear term $\sum_{k=1}^{K} c_{ijk} x_{ijk}$ becomes a piecewise linear function of χ_{ij}, for each (i,j)$\in$A. Figure 4.3 depicts this function. In order to contend with this nondifferentiability, we will utilize the continuous nature of this transformation in the sequel to implement a computationally effective search procedure, assuming a close smooth approximation to the actual function.

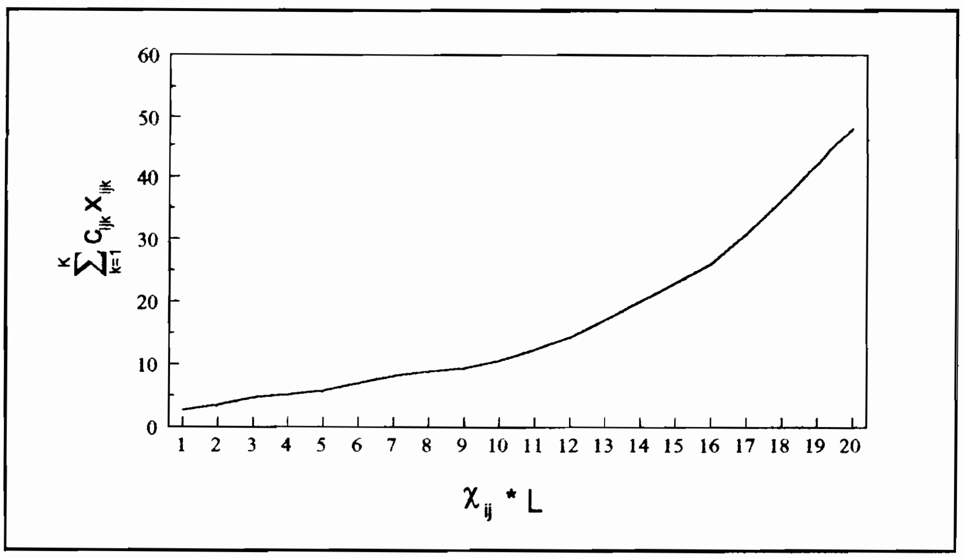

Figure 4.3 NOP objective function terms as a function of χ_{ij}

To summarize, by fixing the X variables *a priori* in the reliability and cost model and by introducing the composite χ variables, we have simplified PNOP to a search process over a simple lower and upper bounded feasible region with no constraints. The χ variables will be used in the higher level conjugate gradient search algorithm, while the inverse-transformed x variables will be used in a lower level line search procedure that utilizes the subproblems SP.

4.3 The Projected Network Optimization Decomposition Algorithm

In this section we present a procedure for solving the decomposed, projected problem PNOP. A conjugate gradient search strategy is implemented in the space of the composite ω variables. At each point in the search, these variables ω are transformed back into the original w variables and the subproblem SP(w) is solved using the algorithm of Section 4.4. The resulting artificial flows from SP(w) are then used to penalize the cost function, and a new search direction is generated. The process is repeated until a solution is reached that has zero artificial flows for which no further improvements are possible. As with all existing procedures for this class of problems, no claim of global optimality is made. However, as we shall see in the computational results section, the structure of the problem and the nature of the search process enables us to identify better solutions than competing methods.

The procedure is initialized at some solution $w^0 \epsilon W$ (not necessarily feasible to NOP). At any iteration $k \geq 0$, given $w^k \epsilon W$, the subproblem $SP(w^k)$ is solved to produce the flow variables $(Q^k,y^k) \equiv (Q^*(w^k),y^*(w^k))$ that satisfy the Hazen-Williams head loss equations, the flow continuity equations, and the associated hydraulic head and flow requirements. If w^k is feasible to NOP, then the artificial flow variables y^k will be zero (Theorem 2) and no penalty will be accrued in PNOP. The objective function will then simply represent the cost of implementing the design w^k. On the other hand, if w^k is not feasible to NOP, then some y^k variables will be positive in the solution to problem $SP(w^k)$, and this will result in a penalty being incurred in the objective function of problem PNOP via the term associated with the parameter μ. Next, the process computes an approximate gradient at ω^k in the transformed composite variable space, where $\omega \equiv (\chi,H_s)$, utilizing a forward finite differencing technique. (We remark here that strictly speaking, the objective function is nondifferentiable, and so, the finite difference gradient estimates are being obtained for some smooth approximation to this function. Furthermore, we can alternatively estimate this gradient analytically instead of using finite differences by adopting a barrier function formulation of SP - see Sections 4.5 and 4.6 below.) Based on the antigradient, or its deflection in a conjugate direction type of approach, a search direction is determined and an inexact line search is performed along this direction in the (composite) design variable space to find a corresponding improved solution w^{k+1}. The inexact line searches are performed based on the quadratic fit line search procedure (see Bazaraa, Sherali, and Shetty (1993), for example), but are not guaranteed to convergence to an optimal solution along the direction of search since the PNOP objective function is not strictly quasi-convex. However, we are only looking for good local optimal solutions to PNOP. Moreover, even under convexity conditions, these inexact line searches provide a computationally effective strategy, and are invariably used. In this context, "improved" solutions are defined by the objective function of PNOP as having a lower combined cost of construction ($cx+CX+c_sH_s$) plus the penalty cost $\mu[y^*_{p0}(w) + y^*_{0q}(w)]$. This total function value is evaluated in the line search process for each tested setting of the composite design variables ω by transforming to the original variables w using Eq. (4.12) and solving the corresponding subproblem SP(w).

In this connection, note that if the penalty parameter μ is large enough, the search procedure will attempt to move away from solutions $w \in W$ that require artificial flows, and hence, will tend to move towards solutions $w \in W$ that are feasible to NOP. Once a feasible solution is found, the search procedure naturally attempts to maintain (near) feasibility, while seeking solutions that have lower design costs. However, if μ is made too large while the iterates are remote from optimality, then numerical difficulties will result as the process zigzags on the boundary of the feasible region and the search process can terminate prematurely at undesirable solutions. Hence, the strategy adopted is to progressively increase μ, starting with a relatively smaller value and sequentially updating the solution to PNOP by periodically doubling the value of μ until a desirable solution having zero artificial flows is obtained. This is akin to the strategy suggested by Fiacco and McCormick in their Sequential Unconstrained Minimization Technique (SUMT) (1964a,b). For this procedure to operate robustly, accurate solutions are required to the subproblems SP in order to accurately determine the penalized artificial flows in PNOP, and hence, to obtain reliable gradient estimates.

Computational experience has shown that by adopting a quadratic penalty function $\mu\,[y^*_{p0}(w) + y^*_{0q}(w)]^2$ during the initial iterations of the above procedure, while the solutions are ostensibly far from an optimum, allows the search to be guided toward improved solutions more rapidly. This reduces the possibility of a premature termination at undesirable solutions far from optimality. However, as the artificial flows approach zero, the quadratic penalty function becomes very flat, and zigzagging again becomes a problem as the algorithm cannot appreciably discriminate between small differences in the objective function. We have found that a two phase approach where the algorithm returns to the linear penalty function $\mu\,[y^*_{p0}(w) + y^*_{0q}(w)]$ when the sum $(y^*_{p0}(w) + y^*_{0q}(w))$ decreases below a threshold of 1.0 works well in practice.

It should also be noted that the above procedure does not guarantee that a global optimum solution will be obtained (see Section 4.6 below for a global optimization approach for problem NOP). Theorem 3 only asserts that *if* a global optimum is found with zero artificial flows, then this is a global solution to problem NOP. However, good quality local optima can be obtained when good starting solutions are used. Since the formulation used in this paper is so different from the ones used in Alperovits and Shamir (1977), Quindry, et al. (1979) and Fujiwara, et al. (1987), the local optima resulting from their algorithms tend *not* to be local optima under this formulation. Thus, there is a good chance that by simply starting from the best solution found from such models (or from some heuristic design process), greatly improved solutions can be found.

The detailed algorithm can now be stated as follows.

Algorithm for Problem PNOP

0. Overall Initialization. For a fixed network configuration, let z and X be given. Recall that $Q_R=(Q,y)$ and $w=(\mathbf{x},X,H_s)$. Define $\omega=(\chi,H_s)$, where χ represents the composite $\mathbf{x}$ variables, and compute the upper bounds χ^U on χ given by $\chi^U_{ij} = KL'_{ij}$ for all (i,j), where $L'_{ij} = L_{ij} - \sum_k X_{ijk}$, is defined by Eq. (4.9). (Here, we assume that K=20 standard diameters are available.) Let the initial solution be given by $\omega^1=(\chi^1,H_s^1)=(\mathbf{L}',\mathbf{0})$, where $\mathbf{L}'$ denotes a vector of the terms L'_{ij}, and where $\mathbf{0}$ denotes a vector of 0's, and let $w^1=(x^1,X,H_s^1)$ be the corresponding design vector obtained by applying the transformation (4.12) on ω^1. Solve $SP(w^1)$ to obtain an optimum solution $Q_R{}^*=(Q^*,y^*)\equiv(Q^*(w^1),y^*(w^1))$. Compute an initial

value of μ such that $\mu[y^*_{p0}(w^1) + y^*_{0q}(w^1)] = 2[cx^1 + CX + c_s H_s^1]$. Select a value for the step size tolerance λ_{tol} (recommended value for λ_{tol}=1.0), the gradient tolerance ϵ (recommended value for ϵ is for 100.), and the artificial flows tolerance y_{tol} (recommended value for y_{tol}=1.0). Recall that the algorithm is run twice; first with f_μ defined using a quadratic penalty function (initialize thus), and then with a linear penalty function (see Eq. 4.7a). Set k=1, RESTART=1, and proceed to Step 2.

1. Initialization for a revised μ. Set k=1, and RESTART=1.

2. Solve the barrier subproblem. Apply the inverse transformation to the composite χ^k variables to obtain values x^k for the original x variables. Let $w^k=(x^k,X,H_s^k)$. Solve SP(w^k) yielding $Q_R^*=(Q^*,y^*)\equiv(Q^*(w^k),y^*(w^k))$.

3. Determine the gradient. Compute the projected estimated antigradient $g^k \equiv \text{proj}[-\nabla f_\mu(\omega^k)]$ of the objective function of problem PNOP at ω^k. (See Section 4.6 for a method for computing the gradient estimate $\nabla f_\mu(\omega^k)$.) The proj[·] projection operation into the ω feasible region is performed as follows. The vector g^k is given by $-\nabla f_\mu(\omega^k)$ except that if any component r has $-\nabla f_\mu(\omega^k)_r>0$ and ω^k_r is at its upper bound, or $-\nabla f_\mu(\omega^k)_r<0$ and ω^k_r is at its lower bound, then we set $g^k_r=0$. If $\|g^k\|<\epsilon$, then proceed to Step 6.

4. Determine the conjugate search direction and the maximum step size. If RESTART=1, then let the direction of motion be given by $d^k = g^k$. Else let $d^k = g^k + \alpha_k d^{k-1}$, where the scalar parameter $\alpha_k = \dfrac{q_k{}^t g^k + (1/\lambda_{k-1}) p_k{}^t g^k}{-q_k{}^t d^{k-1}}$ where $q_k \equiv [g^k - g^{k-1}]$ and $p_k \equiv \omega^k - \omega^{k-1} \equiv \lambda_{k-1} d^{k-1}$.

This choice for α_k is derived using the quasi-Newton condition by Sherali and Ulular (1990) and is recommended when employing inexact line searches. Observe that this is a modified form of Hestenes and Stiefel's (1952) formula. Now, compute the maximum step length $\lambda=\lambda_{k\,max}$ for which $\omega_{LB} \le \omega^k+\lambda d^k \le \omega_{UB}$, where ω_{LB} and ω_{UB} are vectors of lower and upper bounds on the ω variables, respectively, based on the bounds on the original design variables.

5. Line Search. Perform an inexact quadratic fit line search (see Bazaraa, Sherali, and Shetty, 1993, for example) along the direction d^k over $[0,\lambda_{k\,max}]$ to find a step length λ_k. Let

$\omega^{k+1} = \omega^k + \lambda_k d^k$. In case $\lambda_k \leq \lambda_{tol}$, then go to Step 6 if RESTART=1, and otherwise, set RESTART=1. Else, if $\lambda_k > \lambda_{tol}$, then set RESTART=0 in case $\lambda_k < \lambda_{max}$, and set RESTART=1 otherwise. Also, if n_{PNOP} consecutive iterations have been performed with RESTART=0, then set RESTART=1, where n_{PNOP} is the dimension of ω. Increment k by 1 and return to Step 2.

6. Convergence check. If $(y_{p0}^* + y_{0q}^*) \leq y_{tol}$, then stop with ω^k as the prescribed solution in case the linear penalty function is being used, or else, switch over to this penalty function in case the quadratic penalty function is being used using the same y_{tol}. In case this switch in penalty functions is made or if $(y_{p0}^* + y_{0q}^*) > y_{tol}$, then set $\mu \leftarrow 2\mu$ and $\omega^1 \leftarrow \omega^k$, and return to Step 1.

4.4 Algorithm for Solving the Subproblem SP

Recall that the subproblem SP(w) as defined by Eqs. (4.2a)-(4.2b) for a fixed $w=(x,X,H_s)$, is required to be solved in order to find the artificial flows y_{p0} and y_{0q} that are penalized in the objective function of problem PNOP. We solve this convex network flow programming problem using a conjugate gradient search scheme that takes full advantage of the network structure of the constraints in order to simplify the required matrix computations. Inexact line searches are performed based on the quadratic fit line search procedure as before. These line searches have guaranteed convergence to optimal solutions since the subproblems SP(w) are convex (see Bazaraa, Sherali, and Shetty, 1993). Termination of the overall algorithm is based on a gradient tolerance check or a failure to improve the objective function by at least 1% over four consecutive iterations. The detailed algorithm can be stated as follows.

0. Overall Initialization. If this is the first call to SP, then pick the initial feasible solution (Q^1,y^1) by setting $Q^1_{ij}=0$ $\forall(i,j)\in A$, $Q^1_{i0}=b_i$ $\forall i\in S$, $y^1_{p0}=0$, $y^1_{0q}=-\sum(b_i)$, and $y^1_{qi}=-b_i$ $\forall i\in D$. Else, let $Q^1=Q^*(w')$ and $y^1=y^*(w')$ where w' is the set of design variables used at the last call to SP.

1. Initialization. Pick an optimality check tolerance ϵ_{tol}=.001. Set k=1, and RESTART=1.

2. Partitioning of the variables into basic and nonbasic variables. Determine the m-1 maximal spanning tree arcs of $Q_R^k=(Q^k,y^k)$ using Kruskal's algorithm (see Bondy and

Murty, 1976). Let these variables be called basic variables, and let the rest of the variables be called nonbasic variables. Denote the conservation of flow equations matrix by **A**, the matrix of basic columns by **B**, and that of the nonbasic columns by **N**. (Note that these bold type **A** and **N** should not be confused with the arc set A and node set N from Chapter 3.)

3. Determine the reduced gradient. Compute the gradient $\nabla F(Q_R^k,w)$ of the objective function of problem SP(w) at $Q_R = Q_R^k$. Compute $v \equiv \nabla_B F(Q_R^k,w)^t \mathbf{B}^{-1}$ by solving $v\mathbf{B} = \nabla_B F(Q_R^k,w)^t$ on the basis tree. Hence, determine the reduced gradient r_N in the space of the nonbasic variables according to $r_N \equiv \nabla_N F(Q_R^k,w)^t - \nabla_B F(Q_R^k,w)^t \mathbf{B}^{-1}\mathbf{N} = \nabla_N F(Q_R^k,w)^t - v\mathbf{N}$.

4. Determine the search direction and the maximum step size. Compute the projected reduced gradient g_N^k as the vector r_N, except that the j^{th} component is set at zero if $(r_N)_j>0$ and the j^{th} nonbasic variable $(Q_R^k)_{N\,j}=0$. (Note that we are interested in $-g_N^k$ or its deflection as a potential search direction.) If $\|g_N^k\| < \epsilon_{tol}$, then stop with Q_R^k as (near) optimal. If RESTART=1, then let $d_N^k \equiv -g_N^k$. Else, let $d_N^k \equiv -g_N^k + \psi_k d_N^{k-1}$ where the scalar parameter $\psi_k = \dfrac{q_k^{\,t} g_N^k - (1/\lambda_{k-1}) p_k^{\,t} g_N^k}{q_k^{\,t} d_N^{k-1}}$ where $q_k \equiv g_N^k - g_N^{k-1}$ and $p_k \equiv (Q_R^k)_N - (Q_R^{k-1})_N \equiv \lambda_{k-1} d_N^{k-1}$. If $(d_N^k)^t r_N \geq -0.8\|g_N^k\|^2$ then put RESTART=1 and let $d_N^k \equiv -g_N^k$. Compute $u \equiv -\mathbf{N}d_N^k$ utilizing the network structure of **N**, and determine $d_B^k = -\mathbf{B}^{-1}\mathbf{N}d_N^k = \mathbf{B}^{-1}u$ by solving $\mathbf{B}d_B = u$ on the basis tree. Determine the maximum step size $\lambda_{k\,max}$ using the minimum ratio test to obtain $\lambda_{k\,max} = \text{minimum}\{\dfrac{(Q_R^k)_j}{-d^k_j} : d^k_j < 0\}$. If $\lambda_{k\,max} = 0$, then this indicates a degenerate pivot; pivot out the blocking basic variable in exchange for an increasing nonbasic variable in the current direction, set RESTART=1, and return to Step 3. (As a cycling check, the procedure may be terminated if a given number of such consecutive degenerate pivots have been performed.)

5. Line search. Perform an inexact quadratic fit line search (see Bazaraa, Sherali, and Shetty, 1993, for example) along the direction d^k over $[0,\lambda_{k\,max}]$ yielding the step length λ_k. Set $Q_R^{k+1} = Q_R^k + \lambda_k d^k$. If the objective function has not improved by at least 1% over the last 4 iterations, then terminate the algorithm. Otherwise, if $\lambda_k = \lambda_{k\,max}$ then set RESTART=1. Also, if n_{SP} iterations have been performed since the last restart, then set RESTART=1. (n_{SP} is the number of variables in SP.) Else, set RESTART=0. After the line

search some pipes may have both $Q^k_{ij}>0$ and $Q^k_{ji}>0$, say $Q^k_{ij} \geq Q^k_{ji}>0$. Eliminate these two-way flows by setting $Q^k_{ij} = Q^k_{ij} - Q^k_{ji}$ and $Q^k_{ji} = 0$ in this assumed case, and similarly handling the case when $Q^k_{ij} < Q^k_{ji}$. Set $k \leftarrow k+1$. Return to Step 2 if RESTART=1 and to step # if RESTART=0.

4.5 Algorithm for Estimating Gradients Via a Barrier Subproblem BSP

In order to compute reliable gradient estimates for the objective function of problem PNOP, we must formulate a barrier subproblem that does not have discontinuous jumps in the Q_R values for arbitrarily small changes in the w variables. These jumps are caused in the current formulation of the subproblem SP since small changes in the objective function as dictated by small perturbations in w can cause the optimal solution flows $y^*_{p0}(w)$ and $y^*_{0q}(w)$ to "jump" from one extreme point to another. This in turn creates a discontinuity in the objective function of problem PNOP.

To simplify notation, let us rewrite problem SP(w) in matrix form as follows.

$$
\begin{aligned}
\text{SP(w):}\quad & \underline{\text{Minimize}} && F(Q_R, w) \\
& \underline{\text{subject to}} && AQ_R = b \\
& && Q_R \geq 0
\end{aligned}
$$

Here, AQ_R=b represents the constraints (4.1a) - (4.1e) with obvious notation. The corresponding barrier problem (see Bazaraa et al., 1993, for example) for an infinitesimal barrier parameter σ>0 is defined as follows.

$$
\begin{aligned}
\text{BSP(w):}\quad & \underline{\text{Minimize}} && F_\sigma(Q_R, w) \equiv F(Q_R, w) - \sigma \sum_j (\ln (Q_R)_j) \\
& \underline{\text{subject to}} && AQ_R = b
\end{aligned}
$$

Note that the term $\sigma \sum_j (\ln (Q_R)_j)$ automatically limits Q_R to be positive when continuous line searches are performed starting with interior (positive) solutions. Furthermore, we can assume that this term is negligible since σ is infinitesimally small. In fact, $\sigma \sum_j (\ln (Q_R)_j) \to 0$ at optimality in BSP(w) as $\sigma \to 0^+$, and the optimal solution to the original problem SP(w) is obtained via the limit of a sequence of optimal solutions to problem $F_\sigma(Q_R, w)$ as $\sigma \to 0^+$ (see Bazaraa et al., 1993). More importantly, the necessary and sufficient KKT conditions for the optimality of problem BSP(w) as a function of w yield optimal solutions as differentiable functions of w as will be seen in the sequel. This

effectively eliminates the discontinuities in problem PNOP without significantly affecting the values of PNOP at points where no discontinuity exists.

The barrier subproblem BSP(w), for a fixed $w=(x,X,H_s)$, is solved as follows.

1. Solve SP(w) as in Section 4.4 above.
2. Perturb the optimal solution $Q_R^k=(Q^k,y^k)$ to a positive feasible solution $Q_R^+=(Q^+,y^+)$ using a small $\epsilon>0$ as follows:
   - (i) Initialize $Q_R^+ = Q_R^k$.
   - (ii) For each $(i,j)\epsilon A$ with $Q_{ij}^k \geq Q_{ji}^k=0$, let $Q_{ij}^+ \leftarrow Q_{ij}^k + \epsilon$ and $Q_{ji}^+ \leftarrow \epsilon$.
   - (iii) For each $i\epsilon D$, let $y_{ip}^+ \leftarrow y_{ip}^k + \epsilon$, $y_{qi}^+ \leftarrow y_{qi}^k + \epsilon$. Correspondingly, let $y_{p0}^+ \leftarrow y_{p0}^k + |D|\epsilon$ and $y_{0q}^+ \leftarrow y_{0q}^k + |D|\epsilon$.
   - (iv) For each $i\epsilon S$ with $Q_{i0}^k=0$, let $Q_{ir}^+ \leftarrow Q_{ir}^+ - \epsilon$, where $Q_{ir}^+ \equiv \max_s\{Q_{is}^k\}$, and let $Q_{i0}^+ \leftarrow \epsilon$, $y_{0q}^+ \leftarrow y_{0q}^+ + \epsilon$, $y_{qr}^+ \leftarrow y_{qr}^+ + \epsilon$.
   - (v) For each $(i,j)\epsilon A$ with $i\epsilon S$, $j\epsilon D$ and $Q_{ij}^k=0$ (but for which Q_{ji} does not exist), let $Q_{ir}^+ \leftarrow Q_{ir}^+ - \epsilon$, where $Q_{ir}^+ \equiv \max_s\{Q_{is}^+\}$, and let $Q_{ij}^+ \leftarrow \epsilon$, $y_{jp}^+ \leftarrow y_{jp}^+ + \epsilon$, $y_{p0}^+ \leftarrow y_{p0}^+ + \epsilon$, $y_{0q}^+ \leftarrow y_{0q}^+ + \epsilon$, $y_{qr}^+ \leftarrow y_{qr}^+ + \epsilon$.
3. Polish Q_R^+ toward an optimal solution for the barrier problem BSP(w) using a reduced gradient conjugate direction search strategy similar to that used in Section 4.4 for solving SP(w).

4.6 Computing Gradients for PNOP

Let us begin by deriving the gradient for the objective function (4.7a) of problem PNOP by differentiating this function $f_\mu(w)$ with respect to each of the x and H_s variables defining w. (The case of the quadratic penalty function can be treated similarly with trivial modifications.) Recall that X is now fixed, and note that we are assuming that $y^*(w)$ is defined as the optimum solution to the barrier problem BSP(w). Following this, we will derive the estimated gradient with respect to the composite variables ω.

Note that the gradient components with respect to the x_{ijk} and $(H_s)_i$ variables, $\forall ijk$, are given by

$$\frac{\partial f_\mu}{\partial x_{ijk}} = c_{ijk} + \mu \left\{ \frac{\partial y^*_{p0}(w)}{\partial x_{ijk}} + \frac{\partial y^*_{0q}(w)}{\partial x_{ijk}} \right\}, \tag{4.13}$$

and

$$\frac{\partial f_\mu}{\partial (H_s)_i} = (c_s)_i + \mu \left\{ \frac{\partial y^*_{p0}(w)}{\partial (H_s)_i} + \frac{\partial y^*_{0q}(w)}{\partial (H_s)_i} \right\}. \tag{4.14}$$

Thus we need to compute the derivatives of $y^*_{p0}(w)$ and $y^*_{0q}(w)$ with respect to the w variables x_{ijk} and $(H_s)_i$ $\forall$ i,j,k.

Since $y^*_{p0}(w)$ and $y^*_{0q}(w)$ are the optimal flows in the barrier subproblem BSP(w), we can determine the partials required in Eqs. (4.13) and (4.14) from the optimality conditions for BSP(w). At optimality for BSP(w), we obtain some $\bar{Q}_R = (\bar{Q}_{R_B}, \bar{Q}_{R_N}) > 0$, where $\bar{Q}_{R_B}$ are the basic variable components which are all sufficiently positive in the case of nondegeneracy, and $\bar{Q}_{R_N}$ are the nonbasic components which may have some values significantly positive and others positive but near zero. Accordingly, let us partition $\mathbf{A}=(\mathbf{B},\mathbf{N})$ where $\mathbf{B}$ is the nonsingular matrix whose columns correspond to the basic variables. From $\mathbf{A}Q_R=b$, we obtain

$$Q_{R_B} = \mathbf{B}^{-1}b - \mathbf{B}^{-1}\mathbf{N}Q_{R_N}. \tag{4.15}$$

Since the objective function $F_\sigma(Q_R, w)$ for BSP(w) is differentiable in Q_R, the reduced gradient of $F_\sigma(Q_R, w)$ will be equal to a zero vector at optimality for BSP(w) (see Bazaraa, et al., 1993). Thus,

$$\mathbf{r}^t \equiv (\mathbf{r}_B^t, \mathbf{r}_N^t) = [\mathbf{0},\ \nabla_N F_\sigma(Q_R,\mathbf{w})^t - \nabla_B F_\sigma(Q_R,\mathbf{w})^t \mathbf{B}^{-1}\mathbf{N}]$$
$$= [\mathbf{0},\ \mathbf{0}]$$

$$\Rightarrow \quad \nabla_N F_\sigma(Q_R,\mathbf{w})^t - \nabla_B F_\sigma(Q_R,\mathbf{w})^t \mathbf{B}^{-1}\mathbf{N} = \mathbf{0}$$

$$\Rightarrow \quad \nabla_N F_\sigma(Q_R,\mathbf{w}) - (\mathbf{B}^{-1}\mathbf{N})^t \nabla_B F_\sigma(Q_R,\mathbf{w}) = \mathbf{0} \tag{4.16}$$

Taking partial differentials in (4.16) with respect to each component of w, denoted w_i for notational simplicity, and dropping all arguments, we have for each i,

$$\begin{aligned} &H_N \frac{\partial Q_{R_N}}{\partial w_i} + H_{NB} \frac{\partial Q_{R_B}}{\partial w_i} + \nabla_{Nw_i} F_\sigma \\ &- (B^{-1}N)^t [H_{BN} \frac{\partial Q_{R_N}}{\partial w_i} + H_B \frac{\partial Q_{R_B}}{\partial w_i} + \nabla_{Bw_i} F_\sigma] \\ &= \mathbf{0}, \end{aligned} \tag{4.17}$$

where $H_N = \nabla_{NN} F_\sigma$, $H_{NB} = \nabla_{NB} F_\sigma$ and $H_B = \nabla_{BB} F_\sigma$, and where $\nabla_{Nw_i} F_\sigma$ and $\nabla_{Bw_i} F_\sigma$ are vectors of partial derivatives of the components of $\nabla_N F_\sigma$ and $\nabla_B F_\sigma$ with respect to w_i. From (4.15) we have

$$\frac{\partial Q_{R_B}}{\partial w_i} = (-\mathbf{B}^{-1}\mathbf{N}) \frac{\partial Q_{R_N}}{\partial w_i}. \tag{4.18}$$

Substituting (4.18) into (4.17) gives

$$\begin{aligned} &H_N \frac{\partial Q_{R_N}}{\partial w_i} + H_{NB} (-\mathbf{B}^{-1}\mathbf{N}) \frac{\partial Q_{R_N}}{\partial w_i} + \nabla_{Nw_i} F_\sigma \\ &- (\mathbf{B}^{-1}\mathbf{N})^t [H_{BN} \frac{\partial Q_{R_N}}{\partial w_i} + H_B (-\mathbf{B}^{-1}\mathbf{N}) \frac{\partial Q_{R_N}}{\partial w_i} + \nabla_{Bw_i} F_\sigma] \\ &= \mathbf{0}, \end{aligned} \tag{4.19}$$

from which we get

$$[H_N - H_{NB}(\mathbf{B}^{-1}\mathbf{N}) - (\mathbf{B}^{-1}\mathbf{N})^t H_{BN} + (\mathbf{B}^{-1}\mathbf{N})^t H_B (\mathbf{B}^{-1}\mathbf{N})] \frac{\partial Q_{R_N}}{\partial w_i} \tag{4.20}$$

$$= - \nabla_{N w_i} F_\sigma + (\mathbf{B}^{-1}\mathbf{N})^t \nabla_{B w_i} F_\sigma .$$

This system of equations is of the form

$$G \frac{\partial Q_{R_N}}{\partial w_i} = - \nabla_{N w_i} F_\sigma + (\mathbf{B}^{-1}\mathbf{N})^t \nabla_{B w_i} F_\sigma , \tag{4.21}$$

where G is a matrix that is invariant with respect to the particular component w_i for which the partial derivative is being computed. Hence, in order to compute the various $\frac{\partial Q_{R_N}}{\partial w_i}$ $\forall i$, which exist is G is nonsingular, the system of equations (4.21) needs to be solved for various right-hand side vectors, dependent on the particular w_i. The matrix G is therefore factored once, with a suitable perturbation being used if singularity is detected, and the various partial derivative vectors $\frac{\partial Q_{R_N}}{\partial w_i}$ $\forall i$ are then determined easily via (4.21). The corresponding partials $\frac{\partial Q_{R_B}}{\partial w_i}$ $\forall i$ are then computed via the following system of equations implied by (4.18), using the basis tree structure of $\mathbf{B}$:

$$\mathbf{B} \frac{\partial Q_{R_B}}{\partial w_i} = -\mathbf{N} \frac{\partial Q_{R_N}}{\partial w_i} \quad \forall i. \tag{4.22}$$

Now that the derivatives of $y^*_{p0}(w)$ and $y^*_{0q}(w)$ with respect to the w variables x and H_s have been found as desired, the gradient of f_μ can be computed from (4.13) and (4.14).

However, since we are solving the problem PNOP in the space of the composite ω variables, we need to find the derivatives

$$\frac{\partial f_\mu}{\partial \chi_{ij}} \approx \sum_k c_{ijk} \frac{\partial \mathbf{x}_{ijk}}{\partial \chi_{ij}} + \mu \left[\frac{\partial y^*_{p0}}{\partial \chi_{ij}} + \frac{\partial y^*_{0q}}{\partial \chi_{ij}} \right] \quad \forall (i,j) \in A$$

$$\frac{\partial f_\mu}{\partial (H_s)_i} = (c_s)_i + \mu \left[\frac{\partial y^*_{p0}}{\partial (H_s)_i} + \frac{\partial y^*_{0q}}{\partial (H_s)_i} \right] \quad \forall i \in S \tag{4.23}$$

where these equation are an approximation since we are using the right-hand derivatives $\frac{\partial^+ x_{ijk}}{\partial \chi_{ij}}$ of the nondifferentiable functions x_{ijk} of χ_{ij} obtained via (4.12). These right-hand derivatives can easily be found to be 0, +1 or -1 as seen by Eq. (4.12) based on the current value of χ_{ij}. Again we need the derivatives of $y^*_{p0}(w)$ and $y^*_{0q}(w)$ with respect to the components ω_i of ω, and we can compute these as we did before in Eqs. (4.21) and (4.22). Specifically, we obtain $\frac{\partial Q_{R_N}}{\partial \omega_i}$ via the equation

$$G \frac{\partial Q_{R_N}}{\partial \omega_i} = - \nabla_{Nw_i} F_\sigma + (\mathbf{B}^{-1}\mathbf{N})^t \nabla_{Bw_i} F_\sigma \quad \forall i \tag{4.24}$$

and substitute these into the following system to compute $\frac{\partial Q_{R_B}}{\partial \omega_i}$.

$$G \frac{\partial Q_{R_B}}{\partial \omega_i} = -\mathbf{N} \frac{\partial Q_{R_N}}{\partial \omega_i} \quad \forall i \tag{4.25}$$

Here, to compute the components of the vectors $-\nabla_{Nw_i} F_\sigma$ and $\nabla_{Bw_i} F_\sigma$ in (4.24), we again use one-sided right-hand derivatives as in (4.23) by way of an approximation. For example,

$$\nabla_{N\chi_{ij}} F_\sigma \approx \sum_k \frac{\partial}{\partial x_{ijk}} (\nabla_N F_\sigma) \frac{\partial^+ x_{ijk}}{\partial \chi_{ij}} \tag{4.26}$$

where $\dfrac{\partial^+ x_{ijk}}{\partial \chi_{ij}}$ $\forall k$ are computed via (4.12), and where other such derivatives are similarly determined. Observe that by using the composite variables, we have reduced the task to solving (4.23) and (4.24) for only the fewer number of components of ω, rather than for all the components of w.

In this manner, therefore, the computation of the gradient estimates can be made more efficient.

5. A GLOBAL OPTIMIZATION APPROACH FOR THE NETWORK DESIGN PROBLEM

In this chapter we propose an alternative method to the decomposition approach presented in Sections 4.1 through 4.6 for solving the pipe network design problem. We develop an algorithm based on the Reformulation-Linearization Technique (RLT) as propounded by Sherali and Adams (1989, 1990) and Sherali and Tuncbilek (1992) by applying this technique directly to Problem NOP presented in Section 3.2.1 in order to construct a tight linear outer-approximation to this problem. This approximation is imbedded within a branch-and-bound search procedure that is proven to converge to a global ϵ-optimal solution in finite time, for any specified accuracy tolerance $\epsilon>0$.

5.1 Introduction to the RLT Methodology

The fundamental idea behind the RLT strategy is to linearize the nonlinear polynomial terms by substituting linear variables in their stead. However, in order to establish the required relationship between the nonlinear and the linearized terms, various valid or implied constraints are generated and added to the resulting relaxed linear program RLTNOP. These RLT constraints, as they are called, serve to tighten the linear programming relaxation, and their design constitutes the principal step of applying the RLT. In addition, the RLT design must be coordinated with an appropriate partitioning scheme that is applied to the variable bounding intervals. This must be done in a manner so that as these intervals are further and further restricted, the corresponding relaxed linear programs become tighter and tighter, inducing an infinite convergence process to an optimal solution to the original nonlinear program NOP. In other words, we must ensure that along *any* infinite branch of the accompanying branch-and-bound scheme, any accumulation point of the corresponding sequence of solutions generated for the linear program relaxations solves Problem NOP.

In our particular implementation of RLT, we will partition the hyperrectangle defined by the initial bounds on the flow variables into smaller and smaller hyperrectangles by splitting some interval $Q_{ijL} \le Q_{ij} \le Q_{ijU}$ at each branch-and-bound node. Thus at each RLT branch-and-bound node t (where t=1,2,3,...) some of the Q_{ij} variables, $(i,j)\in A$, will have more restricted bounds specified by $Q_{ijL_t} \le Q_{ij} \le Q_{ijU_t}$. Note that we can take the initial node bounds to be $Q_{ijL_1} = 0$ and $Q_{ijU_1} = \sum_{i \in S} b_i$, or preferably, deduce tighter bounds than these based on the structure of the network and any *a priori* knowledge of the problem.

Let us begin our discussion of this procedure by presenting an equivalent restatement of some partitioned subproblem of NOP. Any such subproblem corresponds to an arbitrary hyperrectangle, denoted $\Omega = \{Q \mid Q_L \le Q \le Q_U\}$, over which this node subproblem in the branch-and-bound process needs to be analyzed.

NOP(Ω): <u>Minimize</u> $\quad c \cdot x + C \cdot X + c_s \cdot H_s$ **(5.1a)**

<u>subject to</u>

$$\sum_j Q_{ij} - \sum_j Q_{ji} \le b_i \qquad \forall\, i \in S \tag{5.1b}$$

$$\sum_j Q_{ij} - \sum_j Q_{ji} = b_i \qquad \forall\, i \in D \tag{5.1c}$$

$$H_i + E_i \le F_i + H_{si} \qquad \forall\, i \in S \tag{5.1d}$$

$$(H_i + E_i) - (H_j + E_j) \le Q_{ij}^2\, l(\mathbf{x}, X)_{ij} \qquad \forall\, (i,j) \in A, \text{ with equality holding if } Q_{ij} > 0. \tag{5.1e}$$

$$\sum_{k=1}^{K} (x_{ijk} + X_{ijk}) = L_{ij} \qquad \forall (i,j) \in A,\ i<j \tag{5.1f}$$

$$Q_{ij\,L} \le Q_{ij} \le Q_{ij\,U} \qquad \forall (i,j) \in A \tag{5.1g}$$

$$0 \le H_{si} \le H_{si\,U} \qquad \forall\, i \in S \tag{5.1h}$$

$$0 \leq H_i \leq H_{iU} \qquad \forall\, i \in S \tag{5.1i}$$

$$H_{iL} - E_i \leq H_i \leq H_{iU} - E_i \qquad \forall\, i \in D \tag{5.1j}$$

$$x_{ijk} \geq 0\ ,\quad 0 \leq X_{ijk} \leq a_{ijk} \qquad \forall\, (i,j) \in A,\ i<j,\ \forall\, k \tag{5.1k}$$

Here, we have used the notation $l(\mathbf{x},X)_{ij}$ to represent $\sum_{k=1}^{K}\left\{\alpha_{xijk}x_{ijk} + \alpha_{Xijk}X_{ijk}\right\}$, where α_{xijk} and α_{Xijk} are deduced from Eq. (3.3c), H_{siU} is some practical maximum height by which the head at each source node can be increased, and $H_{iU} = F_i + H_{siU} - E_i$ is the corresponding maximum head for source node i.

To simplify notation, we now introduce three mutually exclusive and exhaustive arc subsets of A. The subset A_N is defined to be the set of arcs whose flow direction is NOT as yet determined. Thus for link (i,j) at node t, when $Q_{ijU} > 0$ and $Q_{jiU} > 0$ then we have $(i,j) \in A_N$ and $(j,i) \in A_N$. Note that this also implies that $Q_{ijL} \equiv Q_{jiL} \equiv 0$. The subset A_D is defined to be the set of arcs whose flow direction is DETERMINED, and corresponding to this, is a subset A_Z of arcs for which the flow is known to be ZERO. Thus when $Q_{ijU} = 0$ we have $(i,j) \in A_Z$ and $(j,i) \in A_D$. Likewise, when $Q_{jiU} = 0$ then we have $(j,i) \in A_Z$ and $(i,j) \in A_D$. Arcs in the subset A_Z can effectively be eliminated from the problem formulation for that particular branch-and-bound node. The union $A_N \cup A_D$, designated A_U, is the set of arcs that have a possibility of nonzero flow. (Notice that $A_U = A - A_Z$.) Hence, we specifically need to consider the flows Q_{ij}, for $(i,j) \in A_N \cup A_D$ in the formulation.

We now make a logical assumption that arcs incident to a source node have their direction of flow determined. That is, for each $(i,j) \in A$ with $i \in S$ and j being a demand node or intermediate node, we have $(i,j) \in A_D$ and $(j,i) \in A_z$. This is a reasonable assumption that requires the flows incident at any source to be directed from the source node to a demand (or intermediate) node. Note that in practice, there might exist flows from one source node to another, but in such cases, we assume that the design implies that the direction of flow between the sources is known *a priori*.

When $(i,j) \in A_D$, then the direction is known to be from node i to node j, and we can reduce the corresponding Eq. (5.1e) in the linear program for that node to the following:

$$(H_i+E_i) - (H_j+E_j) = Q_{ij}^2\, l(\mathbf{x},X)_{ij}. \tag{5.2}$$

On the other hand, when $(i,j)\epsilon A_N$, then Q_{ij} is not necessarily greater than zero. In order to avoid quartic polynomial terms that can greatly increase the size of the resulting RLT relaxations, we substitute the following four approximating inequalities for this one equation.

$$\{(H_i+E_i) - (H_j+E_j)\} \le Q_{ij}^2\, l(\mathbf{x},X)_{ij} \tag{5.3a}$$

$$Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \ge \{2\bar{Q}_{ij}\, Q_{ij}^2 - \bar{Q}_{ij}^2 Q_{ij}\}\, l(\mathbf{x},X)_{ij} \quad \text{for } \bar{Q}_{ij} = 0,\ \frac{Q_{ijU}}{2},\ Q_{ijU} \tag{5.3b}$$

Constraints (5.3a) and (5.3b) represent (5.1e) exactly when $Q_{ij} = 0$. However, when $Q_{ij} > 0$, (5.3a) imposes the required constraint as a "less than or equal to" inequality, while the corresponding "greater than or equal to" inequality is only approximated by the constraints (5.3b). This approximation is exact when $Q_{ij}=\bar{Q}_{ij}$ in (5.3b), and hence its fidelity is improved as the bounding interval for Q_{ij} is reduced in size by the partitioning process. Note that these approximating constraints yield a valid relaxation since by the convexity of Q_{ij}^2, (5.1e) implies that

$$\begin{aligned} Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} &= Q_{ij}^3\, l(\mathbf{x},X)_{ij} \\ &\ge [2\bar{Q}_{ij}\, Q_{ij} - \bar{Q}_{ij}^2\}\, Q_{ij}\, l(\mathbf{x},X)_{ij}. \end{aligned} \tag{5.4}$$

To summarize, the constraints (5.1e) can be represented at any branch-and-bound node as follows.

$$(H_i+E_i) - (H_j+E_j) = Q_{ij}^2\, l(\mathbf{x},X)_{ij} \qquad \forall\,(i,j)\in A_D \tag{5.5a}$$

$$\left.\begin{array}{c}
\{(H_i+E_i) - (H_j+E_j)\} \le Q_{ij}^2\, l(\mathbf{x},X)_{ij} \\
Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \ge 0 \\
Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \ge \{Q_{ijU}Q_{ij}^2 - .25\,Q_{ijU}^2 Q_{ij}\}\, l(\mathbf{x},X)_{ij} \\
Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \ge \{2\,Q_{ijU}\,Q_{ij}^2 - Q_{ijU}^2 Q_{ij}\}\, l(\mathbf{x},X)_{ij}
\end{array}\right\} \forall\,(i,j)\in A_N \tag{5.5b}$$

5.2 Reformulation Phase

We now add the following valid or implied nonlinear constraints to the program NOP(Ω) to capture useful relationships between the inherent polynomial terms and the linear variables that will be used to represent them. These constraints are formed by generating the following product constraints that are implied by the original problem constraints.

$$(5.1f) * Q_{ij} \qquad \forall (i,j) \in A_U,\ i<j \tag{5.6a}$$

$$(5.1f) * Q_{ij}^2 \qquad \forall (i,j) \in A_U,\ i<j \tag{5.6b}$$

$$(5.1f) * Q_{ji} \qquad \forall (j,i) \in A_U,\ i<j \tag{5.6c}$$

$$(5.1f) * Q_{ji}^2 \qquad \forall (j,i) \in A_U,\ i<j \tag{5.6d}$$

$$(5.1j) * (Q_{ij} - Q_{ijL}) \quad \text{where (5.1j) is written for node } i, \quad \forall\,(i,j)\in A_N \tag{5.6e}$$

$$(5.1j) * (Q_{ijU} - Q_{ij}) \quad \text{where (5.1j) is written for node } i, \quad \forall\,(i,j)\in A_N \tag{5.6f}$$

$$(5.1j)*(Q_{ij} - Q_{ijL}) \quad \text{where (5.1j) is written for node } j, \quad \forall (i,j) \in A_N \tag{5.6g}$$

$$(5.1j)*(Q_{ijU} - Q_{ij}) \quad \text{where (5.1j) is written for node } j, \quad \forall (i,j) \in A_N \tag{5.6h}$$

$$(5.1k)*(Q_{ij} - Q_{ijL})(Q_{ijU} - Q_{ij}) \qquad \forall (i,j) \in A_U,\ i<j,\ \forall k \tag{5.6i}$$

$$(5.1k)*(Q_{ij} - Q_{ijL})^2 \qquad \forall (i,j) \in A_U,\ i<j,\ \forall k \tag{5.6j}$$

$$(5.1k)*(Q_{ijU} - Q_{ij})^2 \qquad \forall (i,j) \in A_U,\ i<j,\ \forall k \tag{5.6k}$$

$$(5.1k)*(Q_{ji} - Q_{jiL})(Q_{jiU} - Q_{ji}) \qquad \forall (j,i) \in A_U,\ i<j,\ \forall k \tag{5.6l}$$

$$(5.1k)*(Q_{ji} - Q_{jiL})^2 \qquad \forall (j,i) \in A_U,\ i<j,\ \forall k \tag{5.6m}$$

$$(5.1k)*(Q_{jiU} - Q_{ji})^2 \qquad \forall (j,i) \in A_U,\ i<j,\ \forall k \tag{5.6n}$$

$$(Q_{ij} - Q_{ijL})(Q_{ijU} - Q_{ij}) \geq 0 \qquad \forall (i,j) \in A_U \tag{5.6o}$$

$$(Q_{ij} - Q_{ijL})^2 \geq 0 \qquad \forall (i,j) \in A_U \tag{5.6p}$$

$$(Q_{ijU} - Q_{ij})^2 \geq 0 \qquad \forall (i,j) \in A_U \tag{5.6q}$$

Constraints (5.6o)-(5.6q) result from self-products of the bound factors in (5.1g). Furthermore, it might appear that we could have likewise generated the RLT constraints $(5.1k)^*(Q_{ij} - Q_{ijL})$, $(5.1k)^*(Q_{ijU} - Q_{ij})$, $(5.1k)^*(Q_{ji} - Q_{jiL})$ and $(5.1k)^*(Q_{jiU} - Q_{ji})$. However, these constraints can readily be verified to be implied by various surrogates of the constraints (5.6i)-(5.6n). On the other hand, this reformulation can possibly benefit further by the generation of constraints of the type $(5.1c)^*H_i$ $\forall i \in D$ for which all incident arcs are in A_N, and by the constraints of the type $(Q_{ij} - \bar{Q}_{ij})^2 \geq 0$ where $\bar{Q}_{ij} \in (Q_{ijL}, Q_{ijU})$ corresponds to some chosen grid point(s). Possible improvements from these type of constraints will be investigated in future research.

5.3 Linearization Phase

The principal idea behind the RLT linearization phase is to substitute a single new variable for each quadratic or cubic polynomial term involving the product of some of the original problem variables. Hence, solutions that are feasible to the original problem are always feasible to the linear relaxation, but not vice versa. Therefore, we obtain valid lower bounds. In order to formulate such a linear programming relaxation of NOP, we make the following substitutions of new variables to represent the corresponding nonlinear terms.

$$
\begin{aligned}
&Q_{ij}^2 \rightarrow q_{ij} \\
&Q_{ij}H_i \rightarrow h_{ij}^+ \\
&Q_{ij}H_j \rightarrow h_{ij}^- \\
&Q_{ij}^2 x_{(ij)k} \rightarrow \lambda_{ijk}^+ \\
&Q_{ij} x_{(ij)k} \rightarrow \lambda_{ijk}^- \\
&Q_{ij}^2 X_{(ij)k} \rightarrow \Lambda_{ijk}^+ \\
&Q_{ij} X_{(ij)k} \rightarrow \Lambda_{ijk}^-
\end{aligned}
\tag{5.7}
$$

where $x_{(ij)k} \equiv x_{ijk}$ if i<j and $x_{(ij)k} \equiv x_{jik}$ if j<i. Correspondingly, $X_{(ij)k} \equiv X_{ijk}$ if i<j and $X_{(ij)k} \equiv X_{jik}$ if j<i. The h_{ij}^+ and h_{ij}^- variables are only included in the formulation when the flow direction is not known; that is, when $(i,j) \in A_N$, so that constraints (5.5b) become necessary. The remaining linearized variables are included whenever there is a possibility of a positive flow for the corresponding arc; that is, whenever $(i,j) \in A_U$. Notice that if the relationships in (5.7) hold as an equality for *all* of the RLT variables, then the LP relaxation solves the original problem exactly.

The resulting RLT linear program RLTNOP for a single branch-and-bound node defined by $\Omega = \{Q_{ij} \mid Q_{ijL} \le Q_{ij} \le Q_{ijU} \; \forall (i,j) \in A\}$ can then be stated as follows.

$$\text{RLTNOP}(\Omega): \quad \underline{\text{Minimize}} \quad c\cdot \mathbf{x} + C\cdot X + c_s \cdot H_s \tag{5.8.1}$$

subject to

$$\sum_j Q_{ij} - \sum_j Q_{ji} \le b_i \qquad \forall\, i \in S \tag{5.8.2}$$

$$\sum_j Q_{ij} - \sum_j Q_{ji} = b_i \qquad \forall\, i \in D \tag{5.8.3}$$

$$H_i + E_i \le F_i + H_{si} \qquad \forall\, i \in S \tag{5.8.4}$$

$$\sum_{k=1}^{K} (x_{ijk} + X_{ijk}) = L_{ij} \qquad \forall (i,j) \in A_U,\ i<j \tag{5.8.5}$$

$$(H_i + E_i) - (H_j + E_j) = \sum_{k=1}^{K} \left\{ \alpha_{xijk} \lambda_{ijk}^{+} + \alpha_{Xijk} \Lambda_{ijk}^{+} \right\} \qquad \forall\, (i,j) \in A_D \tag{5.8.6}$$

$$(H_i + E_i) - (H_j + E_j) \le \sum_{k=1}^{K} \left\{ \alpha_{xijk} \lambda_{ijk}^{+} + \alpha_{Xijk} \Lambda_{ijk}^{+} \right\} \qquad \forall\, (i,j) \in A_N \tag{5.8.7}$$

$$h_{ij}^{+} + E_i Q_{ij} - h_{ij}^{-} - E_j Q_{ij} \ge 0 \qquad \forall\, (i,j) \in A_N \tag{5.8.8}$$

$$h_{ij}^{+} + E_i Q_{ij} - h_{ij}^{-} - E_j Q_{ij} \ge \sum_{k=1}^{K} \left\{ Q_{ijU} (\alpha_{xijk} \lambda_{ijk}^{+} + \alpha_{Xijk} \Lambda_{ijk}^{+}) - (Q_{ijU}/2)^2 (\alpha_{xijk} \lambda_{ijk}^{-} + \alpha_{Xijk} \Lambda_{ijk}^{-}) \right\} \qquad \forall\, (i,j) \in A_N \tag{5.8.9}$$

$$h_{ij}^{+} + E_i Q_{ij} - h_{ij}^{-} - E_j Q_{ij} \ge \sum_{k=1}^{K} \left\{ 2\, Q_{ijU} (\alpha_{xijk} \lambda_{ijk}^{+} + \alpha_{Xijk} \Lambda_{ijk}^{+}) - Q_{ijU}^{2} (\alpha_{xijk} \lambda_{ijk}^{-} + \alpha_{Xijk} \Lambda_{ijk}^{-}) \right\} \qquad \forall\, (i,j) \in A_N \tag{5.8.10}$$

$$\sum_{k=1}^{K}(\lambda_{ijk}^{-} + \Lambda_{ijk}^{-}) - L_{ij} Q_{ij} = 0 \qquad \forall (i,j) \in A_U,\ i<j \qquad (5.8.11)$$

$$\sum_{k=1}^{K}(\lambda_{ijk}^{+} + \Lambda_{ijk}^{+}) - L_{ij} q_{ij} = 0 \qquad \forall (i,j) \in A_U,\ i<j \qquad (5.8.12)$$

$$\sum_{k=1}^{K}(\lambda_{jik}^{-} + \Lambda_{jik}^{-}) - L_{ij} Q_{ji} = 0 \qquad \forall (j,i) \in A_U,\ i<j \qquad (5.8.13)$$

$$\sum_{k=1}^{K}(\lambda_{jik}^{+} + \Lambda_{jik}^{+}) - L_{ij} q_{ji} = 0 \qquad \forall (j,i) \in A_U,\ i<j \qquad (5.8.14)$$

$$h_{ij}^{+} - (H_{iL} - E_i)Q_{ij} - Q_{ijL}H_i \geq -(H_{iL} - E_i)Q_{ijL} \qquad \forall (i,j) \in A_N \qquad (5.8.15)$$

$$h_{ij}^{+} - (H_{iU} - E_i)Q_{ij} - Q_{ijL}H_i \leq -(H_{iU} - E_i)Q_{ijL} \qquad \forall (i,j) \in A_N \qquad (5.8.16)$$

$$-h_{ij}^{+} + (H_{iL} - E_i)Q_{ij} + Q_{ijU}H_i \geq (H_{iL} - E_i)Q_{ijU} \qquad \forall (i,j) \in A_N \qquad (5.8.17)$$

$$-h_{ij}^{+} + (H_{iU} - E_i)Q_{ij} + Q_{ijU}H_i \leq (H_{iU} - E_i)Q_{ijU} \qquad \forall (i,j) \in A_N \qquad (5.8.18)$$

$$h_{ij}^{-} - (H_{jL} - E_j)Q_{ij} - Q_{ijL}H_j \geq -(H_{jL} - E_j)Q_{ijL} \qquad \forall (i,j) \in A_N \qquad (5.8.19)$$

$$h_{ij}^{-} - (H_{jU} - E_j)Q_{ij} - Q_{ijL}H_j \leq -(H_{jU} - E_j)Q_{ijL} \qquad \forall (i,j) \in A_N \qquad (5.8.20)$$

$$-h_{ij}^{-} + (H_{jL} - E_j)Q_{ij} + Q_{ijU}H_j \geq (H_{jL} - E_j)Q_{ijU} \qquad \forall (i,j) \in A_N \qquad (5.8.21)$$

$$-h_{ij}^{-} + (H_{jU} - E_j)Q_{ij} + Q_{ijU}H_j \leq (H_{jU} - E_j)Q_{ijU} \qquad \forall (i,j) \in A_N \qquad (5.8.22)$$

$$-\lambda_{ijk}^{+} + (Q_{ijL}+Q_{ijU})\lambda_{ijk}^{-} - Q_{ijL}Q_{ijU}x_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.23)$$

$$-\Lambda_{ijk}^{+} + (Q_{ijL}+Q_{ijU})\Lambda_{ijk}^{-} - Q_{ijL}Q_{ijU}X_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.24)$$

$$\begin{aligned}&-\Lambda_{ijk}^{+} + (Q_{ijL}+Q_{ijU})\Lambda_{ijk}^{-} - Q_{ijL}Q_{ijU}X_{ijk}\\ &\qquad + a_{ijk}q_{ij} - a_{ijk}(Q_{ijL}+Q_{ijU})Q_{ij} \leq -a_{ijk}Q_{ijL}Q_{ijU} \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.25)\end{aligned}$$

$$\lambda_{ijk}^{+} - 2Q_{ijL}\lambda_{ijk}^{-} + Q_{ijL}^{2}x_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.26)$$

$$\Lambda_{ijk}^{+} - 2Q_{ijL}\Lambda_{ijk}^{-} + Q_{ijL}^{2}X_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.27)$$

$$\begin{aligned}&\Lambda_{ijk}^{+} - 2Q_{ijL}\Lambda_{ijk}^{-} + Q_{ijL}^{2}X_{ijk}\\ &\qquad - a_{ijk}q_{ij} + 2a_{ijk}Q_{ijL}Q_{ij} \leq a_{ijk}Q_{ijL}^{2} \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.28)\end{aligned}$$

$$\lambda_{ijk}^{+} - 2Q_{ijU}\lambda_{ijk}^{-} + Q_{ijU}^{2}x_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.29)$$

$$\Lambda_{ijk}^{+} - 2Q_{ijU}\Lambda_{ijk}^{-} + Q_{ijU}^{2}X_{ijk} \geq 0 \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.30)$$

$$\begin{aligned}&\Lambda_{ijk}^{+} - 2Q_{ijU}\Lambda_{ijk}^{-} + Q_{ijU}^{2}X_{ijk}\\ &\qquad - a_{ijk}q_{ij} + 2a_{ijk}Q_{ijU}Q_{ij} \leq a_{ijk}Q_{ijU}^{2} \qquad \forall (i,j)\in A_U,\ i<j,\ \forall k \qquad (5.8.31)\end{aligned}$$

$$-\lambda_{jik}^{+} + (Q_{jiL}+Q_{jiU})\lambda_{jik}^{-} - Q_{jiL}Q_{jiU}x_{ijk} \geq 0 \qquad \forall (j,i)\in A_U,\ i<j,\ \forall k \qquad (5.8.32)$$

$$-\Lambda_{jik}^{+} + (Q_{jiL}+Q_{jiU})\Lambda_{jik}^{-} - Q_{jiL}Q_{jiU}X_{jik} \geq 0 \qquad \forall (j,i)\in A_U,\ i<j,\ \forall k \qquad (5.8.33)$$

$$\begin{aligned}&-\Lambda_{jik}^{+} + (Q_{jiL}+Q_{jiU})\Lambda_{jik}^{-} - Q_{jiL}Q_{jiU}X_{ijk}\\ &\qquad + a_{ijk}q_{ji} - a_{ijk}(Q_{jiL}+Q_{jiU})Q_{ji} \leq a_{ijk}Q_{jiL}Q_{jiU} \qquad \forall (j,i)\in A_U,\ i<j,\ \forall k \qquad (5.8.34)\end{aligned}$$

$$\lambda_{jik}^{+} - 2Q_{jiL}\lambda_{jik}^{-} + Q_{jiL}^{2}x_{ijk} \geq 0 \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.35)$$

$$\Lambda_{jik}^{+} - 2Q_{jiL}\Lambda_{jik}^{-} + Q_{jiL}^{2}X_{ijk} \geq 0 \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.36)$$

$$\Lambda_{jik}^{+} - 2Q_{jiL}\Lambda_{jik}^{-} + Q_{jiL}^{2}X_{ijk}$$
$$-a_{ijk}q_{ji} + 2a_{ijk}Q_{jiL}Q_{ji} \leq a_{ijk}Q_{jiL}^{2} \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.37)$$

$$\lambda_{jik}^{+} - 2Q_{jiU}\lambda_{jik}^{-} + Q_{jiU}^{2}x_{ijk} \geq 0 \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.38)$$

$$\Lambda_{jik}^{+} - 2Q_{jiU}\Lambda_{jik}^{-} + Q_{jiU}^{2}X_{ijk} \geq 0 \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.39)$$

$$\Lambda_{jik}^{+} - 2Q_{jiU}\Lambda_{jik}^{-} + Q_{jiU}^{2}X_{ijk}$$
$$-a_{ijk}q_{ji} + 2a_{ijk}Q_{jiU}Q_{ji} \leq a_{ijk}Q_{jiU}^{2} \qquad \forall\,(j,i)\epsilon A_U,\ i<j,\ \forall\,k \qquad (5.8.40)$$

$$-q_{ij} + (Q_{ijU}+Q_{ijL})Q_{ij} - Q_{ijL}Q_{ijU} \geq 0 \qquad \forall(i,j)\in A_U \qquad (5.8.41)$$

$$q_{ij} - 2Q_{ijL}Q_{ij} + Q_{ijL}^{2} \geq 0 \qquad \forall(i,j)\in A_U \qquad (5.8.42)$$

$$q_{ij} - 2Q_{ijU}Q_{ij} + Q_{ijU}^{2} \geq 0 \qquad \forall(i,j)\in A_U \qquad (5.8.43)$$

$$Q_{ij\,L} \leq Q_{ij} \leq Q_{ij\,U} \qquad \forall(i,j)\in A_U \qquad (5.8.44)$$

$$0 \leq H_{si} \leq H_{si\,max} \qquad \forall\,i\epsilon S \qquad (5.8.45)$$

$$0 \leq H_{i} \leq H_{imax} \qquad \forall\,i\epsilon S \qquad (5.8.46)$$

$$H_{iL} - E_i \leq H_i \leq H_{iU} - E_i \qquad \forall\,i\epsilon D \qquad (5.8.47)$$

$$x_{ijk} \geq 0 \qquad \forall\,(i,j) \in A_U,\ i<j,\ \forall\, k \tag{5.8.48}$$

$$0 \leq X_{ijk} \leq a_{ijk} \qquad \forall\,(i,j) \in A_U,\ i<j,\ \forall\, k \tag{5.8.49}$$

$$Q_{ijL}{}^2 \leq q_{ij} \leq Q_{ijU}{}^2 \qquad \forall (i,j) \in A_U \tag{5.8.50}$$

$$Q_{ijL}(H_{iL} - E_i) \leq h_{ij}^{+} \leq Q_{ijU}(H_{iU} - E_i) \qquad \forall (i,j) \in A_N \tag{5.8.51}$$

$$Q_{ijL}(H_{jL} - E_j) \leq h_{ij}^{-} \leq Q_{ijU}(H_{jU} - E_j) \qquad \forall (i,j) \in A_N \tag{5.8.52}$$

$$0 \leq \lambda_{ijk}^{+} \leq Q_{ijU}{}^2 L_{ij} \qquad \forall (i,j) \in A_U,\ \forall k \tag{5.8.53}$$

$$0 \leq \lambda_{ijk}^{-} \leq Q_{ijU} L_{ij} \qquad \forall (i,j) \in A_U,\ \forall k \tag{5.8.54}$$

$$0 \leq \Lambda_{ijk}^{+} \leq Q_{ijU}{}^2 a_{ijk} \qquad \forall (i,j) \in A_U,\ \forall k \tag{5.8.55}$$

$$0 \leq \Lambda_{ijk}^{-} \leq Q_{ijU} a_{ijk} \qquad \forall (i,j) \in A_U,\ \forall k \tag{5.8.56}$$

Note that the bounds on the new RLT variables given by constraints (5.8.50) - (5.8.56) are implied by the other RLT constraints that have been generated. Their purpose in this formulation, however, is for possible use in some decomposition approach such as a Lagrangian dual optimization procedure to solve this linear programming problem. In the same vein, some of the original variable bounding constraints are implied by the new RLT constraints. For example, (5.1k) is implied by surrogating 2*(5.6i) + (5.6j) + (5.6k) in the linearized form. Again, these simple bounding constraints have been retained for convenience.

Lemma 1 below verifies that RLTNOP is indeed a relaxation of NOP (see Sherali and Tuncbilek (1992) for a similar result applied to general polynomial problems). As before, let us denote NOP(Ω) as the problem NOP restricted to the hyperrectangle Ω that constrains the Q_{ij} flows at some branch-and-bound node, RLTNOP(Ω) as the corresponding

RLT linear program representation, and denote $v[\cdot]$ as the objective function value at optimality of any problem $[\cdot]$.

<u>Lemma 1</u>. For any hyperrectangle Ω of Q_{ij} flows, $v[\text{RLTNOP}(\Omega)] \le v[\text{NOP}(\Omega)]$. Moreover, if the optimal solution $(Q^*, H^*, x^*, X^*, q^*, h^{+*}, h^{-*}, \lambda^{+*}, \lambda^{-*}, \Lambda^{+*}, \Lambda^{-*})$ obtained for RLTNOP(Ω) satisfies Equation (5.7) as an equality for each of the RLT variables, or more simply, if (Q^*, H^*, x^*, X^*) is feasible to NOP(Ω), then (Q^*, H^*, x^*, X^*) solves NOP(Ω).

<u>Proof</u>. For any feasible solution $(\bar{Q}, \bar{H}, \bar{x}, \bar{X})$ to NOP(Ω), there exists a feasible solution $(\bar{Q}, \bar{H}, \bar{x}, \bar{X}, \bar{q}, \bar{h}^+, \bar{h}^-, \bar{\lambda}^+, \bar{\lambda}^-, \bar{\Lambda}^+, \bar{\Lambda}^-)$ to RLTNOP(Ω) with the same objective function value which is constructed using the definitions in (5.7). Hence $v[\text{RLTNOP}(\Omega)] \le v[\text{NOP}(\Omega)]$. Moreover, if (5.7) holds as an equality for each variable to an optimal solution $(Q^*, H^*, x^*, X^*, q^*, h^{+*}, h^{-*}, \lambda^{+*}, \lambda^{-*}, \Lambda^{+*}, \Lambda^{-*})$ for RLTNOP(Ω), then (Q^*, H^*, x^*, X^*) is feasible to NOP(Ω) and $v[\text{RLTNOP}(\Omega)]=c\cdot x^* + C\cdot X^* + c_s\cdot H_s^*$, which equals the value of NOP(Ω) at (Q^*, H^*, x^*, X^*). Hence, (Q^*, H^*, x^*, X^*) solves NOP(Ω). ■

Lemma 2 below shows that if at any feasible solution to RLTNOP(Ω), some flow value Q_{ij} for a particular arc (i,j) at a branch-and-bound node is at one of its bounds in the Ω-hyperrectangle defining the node, then the relationships in (5.7) hold as equalities for that variable.

<u>Lemma 2</u>. Let $(Q, H, x, X, q, h^+, h^-, \lambda^+, \lambda^-, \Lambda^+, \Lambda^-)$ be any feasible solution to RLTNOP(Ω) for some defined hyperrectangle Ω, and suppose that an arc flow Q_{ij} satisfies $Q_{ij} = Q_{ijL}$ or $Q_{ij} = Q_{ijU}$ in this solution. Then the RLT-variable relationships in (5.7) hold as equalities for all the variables that are associated with Q_{ij} in this solution.

<u>Proof</u>.

<u>Case 1</u>: Let $Q_{ij} = Q_{ijL}$. Then (5.8.41) $\rightarrow q_{ij} \le Q_{ijL}^2 = Q_{ij}^2$ and (5.8.42) $\rightarrow q_{ij} \ge Q_{ijL}^2 = Q_{ij}^2$. Thus $q_{ij} = Q_{ij}^2$. Likewise (5.8.15) and (5.8.16) $\rightarrow h_{ij}^+ = Q_{ijL}H_i = Q_{ij}H_i$, and (5.8.19) and (5.8.20) $\rightarrow h_{ij}^- = Q_{ijL}H_j = Q_{ij}H_j$. Next, the surrogation of constraints (5.8.23) and (5.8.26) plus the assumption that $(Q_{ijU} - Q_{ijL}) > 0 \rightarrow \lambda_{ijk}^- \ge Q_{ijL}x_{ijk}$ $\forall k$. Similarly, the surrogation of (5.8.24) and (5.8.27) yields $\Lambda_{ijk}^- \ge Q_{ijL}X_{ijk}$ $\forall k$. But (5.8.11) and (5.8.5) yield $\sum_k(\lambda_{ijk}^- + \Lambda_{ijk}^-) = L_{ij}Q_{ijL} = \sum_k(x_{ijk}Q_{ijL} + X_{ijk}Q_{ijL})$ which implies that $\sum_k[(\lambda_{ijk}^- - x_{ijk}Q_{ijL}) + (\Lambda_{ijk}^- - X_{ijk}Q_{ijL})] = 0$. By the

foregoing argument, since each term in brackets $(\cdot)$ in this last expression is nonnegative, we must have $\lambda_{ijk}^{-} = x_{ijk}Q_{ijL} = Q_{ij}\,x_{ijk}\ \forall k$, as well as $\Lambda_{ijk}^{-} = X_{ijk}Q_{ijL} = Q_{ij}\,X_{ijk}\ \forall k$. Now, using the previous results, (5.8.23) $\rightarrow \lambda_{ijk}^{+} \leq (Q_{ijL}+Q_{ijU})\lambda_{ijk}^{-} - Q_{ijL}Q_{ijU}\,x_{ijk} = (Q_{ijL}+Q_{ijU})(Q_{ijL}x_{ijk}) - Q_{ijL}Q_{ijU}\,x_{ijk}$ $= Q_{ijL}^{2}\,x_{ijk} = Q_{ij}^{2}\,x_{ijk}$. Moreover, (5.8.26) $\rightarrow \lambda_{ijk}^{+} \geq 2Q_{ijL}\,\lambda_{ijk}^{-} - Q_{ijL}^{2}\,x_{ijk} = 2Q_{ijL}(Q_{ijL}x_{ijk}) - Q_{ijL}^{2}\,x_{ijk} =$ $Q_{ij}^{2}\,x_{ijk}$. Hence, $\lambda_{ijk}^{+} = Q_{ij}^{2}\,x_{ijk}\ \forall k$. The argument for $\Lambda_{ijk}^{+} = Q_{ij}^{2}\,X_{ijk}$ follows similarly via equations (5.8.24) and (5.8.27).

Case 2: Let $Q_{ij} = Q_{ijU}$. The proof for this case follows identically, using a symmetric argument to Case 1. ■

5.4 Branch-and-Bound Strategy

We can now imbed RLTNOP(Ω) in a branch-and-bound procedure to solve Problem NOP globally to within any specified $\epsilon>0$, or percentage, tolerance. This discussion and the section to follow is based on the development in Sherali and Tuncbilek (1992) where general polynomial programs are addressed, but differs in our treatment of constraints (5.1.e). Our approach is to partition the initial hyperrectangle $Q_L \leq Q \leq Q_U$ into smaller and smaller sub-hyperrectangles, each of which is associated with a node of the branch-and-bound tree. Let the sub-hyper-rectangle associated with node t be denoted by Ω^t, where $\Omega^t \equiv \{Q \mid Q_{L_t} \leq Q \leq Q_{U_t}\}$. Then RLTNOP($\Omega^t$) yields a lower bound for the node subproblem NOP(Ω^t). (Note that we will continue to assume that if $Q_{ijL} = Q_{ijU}$ for any $(i,j)\in A$, then Q_{ij} is fixed at this common value in the problem and is no linger treated as a variable in the computations.) In particular, if $(\bar{Q}, \bar{H}, \bar{x}, \bar{X}, \bar{q}, \bar{h}^{+}, \bar{h}^{-}, \bar{\lambda}^{+}, \bar{\lambda}^{-}, \bar{\Lambda}^{+}, \bar{\Lambda}^{-})$ solves RLTNOP(Ω^t) and the corresponding solution $(\bar{Q}, \bar{H}, \bar{x}, \bar{X})$ is feasible to NOP(Ω^t), then by Lemma 1, $(\bar{Q}, \bar{H}, \bar{x}, \bar{X})$ solves NOP(Ω^t), and being feasible to NOP, the value v[NOP(Ω^t)] $\equiv$ v[RLTNOP(Ω^t)] provides an upper bound for the problem NOP. Hence, we have a candidate for possibly updating the incumbent solution (Q^*, H^*, x^*, X^*) and its value v^* for NOP. In any case, if v[RLTNOP(Ω^t)] $\geq v^*$, we can fathom node t. Hence at any stage s of the branch-and-bound algorithm, we have a set of non-fathomed or *active nodes* denoted as T_s. We now select an active node t^* in T_s that has the least objective function value for the corresponding relaxations RLTNOP($\cdot$) (breaking ties arbitrarily). That is, we select $t^* \in \operatorname{argmin}\{v[\text{RLTNOP}(\Omega^t)] : t \in T_s\}$. Next, we partition the hyperrectangle associated with this node t^* into two sub-hyperrectangles based on a *branching variable* selected according to the following partitioning strategy. As we shall see, the selection of a branching variable according to this strategy is not only important from the viewpoint of computational efficiency, but is critical in ensuring the theoretical convergence of the overall procedure as well.

Partitioning Strategy

Let $(\bar{Q}, \bar{H}, \bar{x}, \bar{X})$ be (part of) an optimal solution to RLTNOP(Ω^t). Define

$$\bar{\phi}_{ij} = \bar{Q}_{ij}^2 l(\bar{x},\bar{X})_{ij} \quad \forall (i,j) \in A_U. \tag{5.9}$$

For each $(i,j) \in A_U$, compute the following discrepancy in the critical nonlinear constraint (5.1e):

$$\delta_{ij} = \begin{cases} |(\bar{H}_i + E_i) - (\bar{H}_j + E_j) - \bar{\phi}_{ij}|, & \forall (i,j) \in A_U \text{ such that } \bar{Q}_{ij} > 0 \\ 0, \quad \text{otherwise} & (\forall (i,j) \in A_U \text{ such that } \bar{Q}_{ij} = 0) \end{cases} \tag{5.10}$$

and select

$$(p,q) = \underset{(i,j) \in A_U}{\operatorname{argmax}} \{\delta_{ij}\}. \tag{5.11}$$

The following lemma asserts a simple, but crucial, fact that will be useful in establishing the convergence of the proposed algorithm.

Lemma 3. With (p,q) selected as in (5.11), if $\delta_{pq}=0$, then the (partial) optimum solution $(\bar{Q}, \bar{H}, \bar{x}, \bar{X})$ to RLTNOP(Ω^t) solves NOP(Ω^t) yielding the same objective value. Hence, the corresponding node t can be fathomed, after updating the incumbent feasible solution, if necessary.

Proof. Let (p,q) be selected by (5.11) and suppose that $\delta_{pq}=0$. Since δ_{pq} is the *maximum* nonlinear discrepancy $|(\bar{H}_i + E_i) - (\bar{H}_j + E_j) - \bar{\phi}_{ij}|$, we therefore have that for *all* arcs with $\bar{Q}_{ij}>0$, $(\bar{H}_i + E_i) - (\bar{H}_j + E_j) = \bar{\phi}_{ij}$. On the other hand, if $\bar{Q}_{ij}=0$, then by (5.7) and Lemma 2, we have $\bar{\lambda}_{ijk} = \bar{\Lambda}_{ijk} = 0\ \forall k$, and so by (5.8.6) and (5.8.7), we have $(\bar{H}_i + E_i) - (\bar{H}_j + E_j) \le 0$. Therefore, constraints (5.1e) are satisfied $\forall(i,j)\in A$ for Problem NOP(Ω^t). All the remaining constraints in NOP(Ω^t) are satisfied explicitly by inclusion into the formulation of Problem RLTNOP(Ω^t). Hence, the optimal solution to RLTNOP(Ω^t) is feasible to NOP(Ω^t) and, by Lemma 1, solves NOP(Ω^t). ■

Having selected (p,q) as in (5.11), if $\delta_{pq}>0$, then we partition the interval $[Q_{ijL}, Q_{ijU}]$ into the two intervals $[Q_{ijL}, \bar{Q}_{ij}]$ and $[\bar{Q}_{ij}, Q_{ijU}]$ at the linear program solution $\bar{Q}_{ij}$. A formal statement of the RLT branch-and-bound algorithm is given below.

Branch-and-Bound Algorithm

0. Initialization Step. Initialize the incumbent solution (Q^*, H^*, x^*, X^*) and v^* based on any known operating point obtained via some local optimization or heuristic process. Set the stage counter s=1, and let $T_1=\{1\}$. Denote $\Omega^{s,t} = \Omega^{1,1} \equiv \Omega$ as the initial hyperrectangle for NOP. Solve RLTNOP($\Omega^{1,1}$) to obtain an optimal solution ($\bar{Q}$, $\bar{H}$, $\bar{x}$, $\bar{X}$, $\bar{q}$, $\bar{h}^+$, $\bar{h}^-$, $\bar{\lambda}^+$, $\bar{\lambda}^-$, $\bar{\Lambda}^+$, $\bar{\Lambda}^-$) of objective function value $LB_{1,1} \equiv v[\text{RLTNOP}(\Omega)]$, and determine a branching variable by using (5.11). If $\delta_{pq} = 0$, or if $v^* = LB_{1,1}$, then stop; by Lemma 3, this solution solves the original problem NOP(Ω). Otherwise, set $t^*=1$, and proceed to Step 1.

1. Partitioning Step (Stage s, s≥1). Having the active node (s,t^*) to be partitioned, and given the choice (p,q) for the branching variable as determined by (5.11), partition this node into sub-nodes as follows. Since $\delta_{pq} > 0$, by Lemma 2,

$Q_{pqL}^{s,t^*} < \bar{Q}_{pq}^{s,t^*} < Q_{pqU}^{s,t^*}$, where $[Q_{pqL}^{s,t^*}, Q_{pqU}^{s,t^*}]$ are the bounds on Q_{pq} in the hyperrectangle Ω^{s,t^*}. Accordingly, partition the set Ω^{s,t^*} into two sub-hyper-rectangles Ω^{s,t_1} and Ω^{s,t_2} that are the same as $\Omega^{s,t}$ except that the restriction on Q_{pq} is given by

$Q_{pq} \in [Q_{pqL}{}^{s,t^*}, \bar{Q}_{pq}{}^{s,t^*}]$ in Ω^{s,t_1}, and

$Q_{pq} \in [\bar{Q}_{pq}{}^{s,t^*}, Q_{pqU}{}^{s,t^*}]$ in Ω^{s,t_2}.

After setting $T_s = (T_s - \{t^*\}) \cup \{t_1, t_2\}$, proceed to Step 2.

2. Bounding Step. Solve the linear program RLTNOP(Ω^{s,t_1}). If this problem is infeasible, then fathom the corresponding node at Step 3. Otherwise, find an optimal solution $(Q, H, x, X)^{s,t_1}$ and denote its objective function value by $LB_{s,t_1} = v$[RLTNOP(Ω^{s,t_1})]. Using this optimal solution in (5.11), determine the corresponding branching variable Q_{pq}. If $\delta_{pq} = 0$, then by Lemma 3, this solution solves the node subproblem NOP(Ω^{s,t_1}). In this case, if $v^* > LB_{s,t_1} \equiv v$[RLTNOP($\Omega^{s,t_1}$)], then update the incumbent solution (Q^*, H^*, x^*, X^*) and v^* accordingly. Else, we have $\delta_{pq} > 0$, and so, store the branching variable index (p,q) to be possibly used later. Repeat Step 2 after replacing t_1 by t_2, and then proceed to Step 3.

3. Fathoming Step. Fathom any nonimproving nodes by setting $T_{s+1} = T_s - \{t \in T_s: LB_{s,t} \geq v^*\}$. If $T_{s+1} = \emptyset$, then stop. Otherwise, update $\Omega^{s+1,t} = \Omega^{s,t}$ and $LB_{s+1,t} = LB_{s,t}$ for all $t \in T_{s+1}$. Increment s by 1, and proceed to Step 4.

4. Node Selection Step. Select an active node (s,t^*), where $t^* \in \operatorname{argmin}\{LB_{s,t} : t \in T_s\}$ is associated with the least lower bound $LB_s \equiv LB_{s,t^*}$ over the active nodes at stage s. Return to Step 1.

5.5 Convergence of the RLT Branch-and-Bound Algorithm

In the spirit of Sherali and Tuncbilek (1992), the following theorem establishes the convergence of the RLT based algorithm for solving the water distribution system network optimization problem.

Theorem 4 (Convergence Result). The above Branch-and-Bound algorithm either terminates finitely with the incumbent solution being optimal to NOP(Ω), or else an infinite sequence of stages is generated. In the latter case, along any infinite branch of the branch-and-bound tree, any accumulation point of the sequence of solutions (Q, H, x, X) generated via the optimal linear programming solutions (Q, H, x, X, q, h^+, h^-, λ^+, λ^-, Λ^+, Λ^-) obtained for the relaxations RLTNOP(Ω) corresponding to the nodes of this branch, solves NOP(Ω).

Proof. The case of finite termination is clear. Hence, suppose that an infinite sequence of stages is generated. Consider any infinite branch of the branch-and-bound tree. Denote the associated nested sequence of partitions as $\{\Omega^{s,t(s)}\}_S$ where the indices (s,t(s)) are used to represent the nodes along this branch, with S being the index set of the corresponding stages at which the nodes were selected for partitioning, so that

$$t(s) \equiv t^* \in \operatorname{argmin} \{v[\text{RLTNOP}(\Omega_{s,t})]: t \in T_s\} \text{ for each } s \in S.$$

Hence, we also have that

$$LB_s = LB_{s,t(s)} \equiv v[\text{RLTNOP}(\Omega^{s,t(s)})] \leq v[\text{RLTNOP}(\Omega^{s,t})] \quad \forall t \in T_s, \text{ for each } s \in S. \tag{5.12}$$

Noting that all the iterates are generated within a bounded set, let $\{(Q, H, x, X, q, h^+, h^-, \lambda^+, \lambda^-, \Lambda^+, \Lambda^-)^{s,t(s)}\}_{S_1}$ be any convergent subsequence of the sequence of linear programming optimal solutions corresponding to the nodes (s,t(s)), $s \in S$, that converges to some accumulation point $(\bar{Q}, \bar{H}, \bar{x}, \bar{X}, \bar{q}, \bar{h}^+, \bar{h}^-, \bar{\lambda}^+, \bar{\lambda}^-, \bar{\Lambda}^+, \bar{\Lambda}^-)$. Let $\{\Omega^{s,t(s)}\}_{S_1}$ denote the corresponding subsequence of hyperrectangular partitions. We must show that $(\bar{Q}, \bar{H}, \bar{x}, \bar{X},)$ solves NOP(Ω).

Note that since $LB_{s,t(s)}$ is the least lower bound at stage s, we have,

$$v[\text{NOP}(\Omega)] \geq LB_{s,t(s)} = v[\text{RLTNOP}(\Omega^{s,t(s)}] \quad \forall s \in S_1.$$

Next, observe that over the infinite subsequence of nodes $\Omega^{s,t(s)}$, $s \in S_1$, there exists a variable Q_{uv} that is branched on infinitely often via the choice (5.11). Let $S_2 \subset S_1$ be the index set corresponding to the nodes where the branching is on Q_{uv} with $\{[Q_{uvL}^{s,t(s)}, Q_{uvU}^{s,t(s)}]\}_{S_2}$ being the corresponding (sub)sequences of interval endpoints.

Now, since Step 1 of the branch-and-bound algorithm enforces Q_{uv} to be partitioned at an interior point to the previous interval, we have that $\{Q_{uvL}^{s,t(s)}\}_{S_2}$ is a monotonically nondecreasing sequence having Q_{uvU}^{1} as an upper bound and that $\{Q_{uvU}^{s,t(s)}\}_{S_2}$ is a monotonically nonincreasing sequence having Q_{uvL}^{1} as a lower bound. Hence these sequences converge to fixed endpoints with the corresponding interval converging to say, $\{[Q_{uvL}^{s,t(s)}, Q_{uvU}^{s,t(s)}]\}_{S_2} \to [\bar{Q}_{uvL}, \bar{Q}_{uvU}]$. Likewise, $\{[Q_{ijL}^{s,t(s)}, Q_{ijU}^{s,t(s)}]\}_{S_2} \to [\bar{Q}_{ijL}, \bar{Q}_{ijU}]$ $\forall (i,j) \in A$. Denote $\bar{\Omega} = [\bar{Q}_L, \bar{Q}_U]$ as the corresponding limiting hyperrectangle thus produced. Clearly, the accumulation point $\bar{Q}$ has coordinate $\bar{Q}_{uv} \in [\bar{Q}_{uvL}, \bar{Q}_{uvU}]$. We will now show that $\bar{Q}_{uv}$ is equal to either $\bar{Q}_{uvL}$ or $\bar{Q}_{uvU}$.

Suppose by contradiction that $\bar{Q}_{uv} \in (\bar{Q}_{uvL}, \bar{Q}_{uvU})$, that is, suppose that $\bar{Q}_{uv}$ is an interior point in this interval. Since $(\bar{Q}, \bar{H}, \bar{x}, \bar{X},)$ is an accumulation point, this means that there must exist an infinite number of points in the sequence $\{(Q, H, x, X)^{s,t(s)}\}_{S_2}$ for which the coordinate Q_{uv} also lies in the interior of the interval $[\bar{Q}_{uvL}, \bar{Q}_{uvU}]$. This contradicts our partitioning methodology in Step 1, since each LP optimal solution in the sequence is used as an endpoint in the next interval, and therefore cannot turn out to be an interior point in the final limiting convergence interval $[\bar{Q}_{uvL}, \bar{Q}_{uvU}]$. Hence, $\bar{Q}_{uv}$ is equal to either $\bar{Q}_{uvL}$ or $\bar{Q}_{uvU}$.

Now, by Lemma 2 and the nature of the partitioning strategy (5.10)-(5.11), we have $\delta_{uv} \to 0$. Since the branching rule in Step 1 selects the flow variable that yields the highest discrepancy in the nonlinear constraints by (5.10), and for the nodes (s,t(s)), $s \in S_2$, this variable is (u,v), we deduce that the nonlinear discrepancies δ_{ij} in (5.10) approach zero for all the flow variables $(i,j) \in A$ as well. By Lemma 3, then, it follows that $(\bar{Q}, \bar{H}, \bar{x}, \bar{X},)$ is feasible to NOP($\bar{\Omega}$) and is therefore feasible to NOP(Ω), since $\bar{\Omega} \subset \Omega$. Therefore, $c \cdot \bar{x} + C \cdot \bar{X} + c_s \cdot \bar{H}_s$ serves as an upper bound for NOP(Ω).

Now, $LB_s \leq v[NOP(\Omega)]\ \forall s \in S_2$. But by (5.12), $\{LB_s\}_{S_2} \equiv \{v[RLTNOP(\Omega^{s,t(s)})]\}_{S_2} \rightarrow c\cdot\bar{x} + C\cdot\bar{X} + c_s\cdot\bar{H}_s$. Hence, $c\cdot\bar{x} + C\cdot\bar{X} + c_s\cdot\bar{H}_s \leq v\,[NOP(\Omega)]$. Thus in the limit, $c\cdot\bar{x} + C\cdot\bar{X} + c_s\cdot\bar{H}_s$ serves as both an upper and lower bound on $v[NOP(\Omega)]$, and so, $(\bar{Q}, \bar{H}, \bar{x}, \bar{X},)$ solves $NOP(\Omega)$. This completes the proof. ■

5.6 Size of the Problem RLTNOP

The size of a particular instance of problem $RLTNOP(\Omega)$ depends on the number of nodes and arcs in the WDS network, on the maximum number of candidate diameters permitted for each link (K), and on the number of links that have their flow direction resolved. The number of constraints is $2|S| + |D| + (9+18K)|A_D| + (16+9K)|A_N|$, where the simple upper and lower bound constraints on the variables are not counted, and where we have used the fact that $|A_U| = |A_D| + |A_N|$ and $|A_U, i<j| = |A_D| + |.5A_N|$. The number of variables is $2|S| + |D| + (2+6K)|A_D| + (4+5K)|A_N|$.

As the number of links in the network that have their direction of flow determined increases, the number of constraints and variables associated with the h^+ and h^- variables decreases. In the worst case, none of the links will have their flow directions determined. In this case, $|A_N| = 2|A|$ and $|A_D| = 0$, and the number of constraints is $2|S| + |D| + (32+18K)|A|$, and the number of variables is $2|S| + |D| + (8+10K)|A|$. In the best case, when all the links have their flow directions determined, $|A_N| = 0$ and $|A_D| = |A|$, and the number of constraints is $2|S| + |D| + (9+18K)|A|$, and the number of variables is $2|S| + |D| + (2+6K)|A|$.

For example, in the test problem where $|S| = 1$, $|D| = 6$, $|A| = 7$ and $K=5$, we obtain 862 constraints and 414 variables in the worst case. In the best case, we obtain 701 constraints and 232 variables. Nearly 20% of the constraints and half the variables are eliminated by resolving the directions of flow.

5.7 Practical Considerations

Notice that if the direction in a link is not as yet determined for a particular branch-and-bound node, then constraints (5.7.7)-(5.7.10), (5.7.15)- (5.7.22) and (5.7.51)-(5.7.52) that contain the variables h_{ij}^+ and h_{ij}^- are formulated for both the arcs associated with this link. However, when the direction is determined in some later node in the branch-and-bound tree, the RLT formulation is greatly simplified since the above mentioned constraints that only approximate the hydraulic equations are replaced by the single exact constraints (5.7.6). Additionally, the h_{ij}^+ and h_{ij}^- variables are eliminated. Furthermore, when a flow Q_{ij} tends to be positive, (5.5a) places a stronger restriction on the relaxed RLT linear programming problem than does (5.5b). As a result, it is useful to determine flow directions for critical links, and if this is not ascertainable *a priori*, then several initial nodes can be thus created in a preprocessing step. Also, when an arc $(p,q) \in A_N$ is selected for branching for the first time, the partition can be implemented at $Q_{pq}=Q_{qp}=0$ forming nodes with sub-hyper-rectangles Ω^{s,t_1} and Ω^{s,t_2} that are the same as $\Omega^{s,t}$ except that for node t_1 we have $Q_{pq} \in [0,\ \bar{Q}_{pqU}^{s,t^*}]$ and $Q_{qp} \in [0,\ 0]$ and for node t_2 we have $Q_{pq} \in [0,\ 0]$ and $Q_{qp} \in [0,\ \bar{Q}_{qpU}^{s,t^*}]$ rather than the standard partition determined above. In this light several nodes can be created in place of the initial node in a preprocessing step that selects partitions at the zero flow point for some critical links in the network.

In practice there may be few or no nodes t that yield solutions to RLTNOP(Ω^t) that turn out to be feasible to NOP(Ω^t) in the branch-and-bound process. Thus the incumbent upper bound will not improve (decrease) early enough to strengthen the fathoming efficiency in the tree. In order to provide for such an improvement, we can attempt to generate feasible solutions to Problem NOP at each node t via some quick heuristic or local optimization process that starts with the LP solution $\bar{Q}$ obtained upon solving RLTNOP(Ω^t). One such heuristic could be the procedure of Alperovits and Shamir (1977).

5.8 Modifications for Rational Exponents

In order to compare results with traditional reference test problems such as the one in Alperovits and Shamir (1977), we can modify the RLT procedure to replace the frictional head loss Equation (3.3c) that we have been using with Eq. (3.3a), which has rational exponents. Thus the constraints (5.1e) become

$$(H_i+E_i) - (H_j+E_j) \leq Q_{ij}^{1.852}\, l(\mathbf{x},X)_{ij} \quad \forall\, (i,j) \in A, \text{ with equality holding if } Q_{ij} > 0 \tag{5.13}$$

with appropriate modifications to $l(\mathbf{x},X)_{ij}$.

As before, we replace these constraints with a set of equivalent inequalities as follows, for all $(i,j)\in A$.

$$\{(H_i+E_i) - (H_j+E_j)\} \leq Q_{ij}{}^{1.852}\, l(\mathbf{x},X)_{ij} \tag{5.14a}$$

$$Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \geq Q_{ij}{}^{2.852}\, l(\mathbf{x},X)_{ij} \tag{5.14b}$$

The extra multiplier of Q_{ij} in (5.14b) enforces the equality in (5.13) when $Q_{ij} > 0$ without violating (5.13) when $Q_{ij} = 0$. Now, we use first-order tangential approximation results for convex and concave functions in order to obtain valid inequalities having integer exponents of order two or less. First, since $Q^{.852}$ is a concave function for $Q \geq 0$, we have for any $\bar{Q}_{ij} \in [Q_{ijL}, Q_{ijU}] - \{0\}$,

$$Q_{ij}{}^{.852} \leq \bar{Q}_{ij}^{.852} + (Q_{ij} - \bar{Q}_{ij})(.852\bar{Q}^{-.148}). \tag{5.15}$$

Multiplying by Q_{ij} yields

$$Q_{ij}{}^{1.852} \leq [.852\bar{Q}^{(-.148)}]Q_{ij}{}^{2} + [.148\bar{Q}_{ij}^{(.852)}]Q_{ij}. \tag{5.16}$$

Likewise using $Q^{1.852}$ as a convex function, for any $\bar{Q}_{ij} \in [Q_{ijL}, Q_{ijU}]$, we obtain

$$Q_{ij}{}^{2.852} \geq [1.852\bar{Q}_{ij}^{(.852)}]Q_{ij}{}^{2} - [.852\bar{Q}_{ij}^{(1.852)}]Q_{ij}. \tag{5.17}$$

Now, we can approximate (5.14a) and (5.14b) via the following valid inequalities, using (5.16) and (5.17), respectively.

$$(H_i+E_i) - (H_j+E_j) \le \{[.852\bar{Q}^{(-.148)}]Q_{ij}^2 + [.148\bar{Q}_{ij}^{(.852)}]Q_{ij}\}\ l(\mathbf{x},X)_{ij} \tag{5.18a}$$

$$Q_{ij}\{(H_i+E_i) - (H_j+E_j)\} \ge \{[1.852\bar{Q}_{ij}^{(.852)}]Q_{ij}^2 - [.852\bar{Q}_{ij}^{(1.852)}]Q_{ij}\}\ l(\mathbf{x},X)_{ij} \tag{5.18b}$$

Note that we can implement (5.18) for several $\bar{Q}_{ij}$'s in $[Q_{ijL}, Q_{ijU}]$, say at $\bar{Q}_{ij} = Q_{ijL}$, $.75Q_{ijL}+ .25Q_{ijU}$, $.5Q_{ijL}+ .5Q_{ijU}$, $.25Q_{ijL}+ .75Q_{ijU}$, and Q_{ijU}. (Naturally, when $Q_{ijL} = 0$, we do not include (5.18a) for $\bar{Q}_{ij} = Q_{ijL}$, since the expression is undefined.) By the nature of the Taylor Series expansion, when an LP optimal value for Q_{ij} corresponds with one of the grid points $\bar{Q}_{ij}$, then the constraints (5.18) ensure that the original constraints (5.14) are satisfied exactly. Therefore the above convergence argument can be seen to apply to this case of rational exponents as well.

Now, however, we need to use these approximating inequalities for all arcs that have a possibility of positive flow ($\forall(i,j)\in A_U$), not just the ones with directions undetermined ($\forall(i,j)\in A_N$). Likewise, the constraints (5.6e)-(5.6h) need to be generated for all $(i,j)\in A_U$ as well, since the h^+ and h^- variables will be present in (5.18) in all such cases. In summary, we retain the formulation (5.8) except that we have (5.8.6)-(5.8.10) replaced with the corresponding linearized RLT constraints derived from (5.18a) and (5.18b), for each $(i,j)\in A_U$, and (5.8.15)-(5.8.22) and (5.8.51)-(5.8.52) are constructed $\forall(i,j)\in A_U$. In addition, since the h_{ij}^+ and h_{ij}^- variables are now present in the formulation whenever Q_{ij} is present, we can easily add the tightening constraints (5.1c)$^*H_i$ (using the same nodes *i* about which the original constraints were generated).

We now proceed in the next chapter to present some computational results using both the integer and rational exponent formulations.

6. COMPUTATIONAL EXPERIENCE ON A STANDARD TEST PROBLEM

6.1 Results From the Decomposition Algorithm of Chapter 4

The decomposition strategy presented in Chapter 4 for solving the network optimization problem is applied to the Alperovits and Shamir (1977) single source test problem. The algorithm is implemented on a 486DX/33MHz IBM PC compatible computer using the DOS operating system. The computer program is written using Microsoft FORTRAN. Finite difference methods were used to compute gradients in RNOP. These results can be found in a previous paper (Sherali and Smith, 1993) and were presented at the International Computer Applications for Water Supply and Distribution conference at De Montfort University, Leicester, England in September, 1993.

The test problem is repeated here as represented by Tables 6.1a-6.1b and Figure 6.1 for convenience. (Note that we have reversed the direction of flow for arc 7 from the convention adopted in Alperovits and Shamir (1977), so that all arcs will have "forward" flow in the direction of increasing node numbers, i.e., the "forward" flow here is *from* node 5 *to* node 7.)

Two cases of pipe diameter restrictions using this algorithm were tested. If the set of candidate pipe diameters are restricted to the ones given in the original paper, the solution presented in Table 6.3 below is obtained using our algorithm. This solutions has a total cost of $441,674, a reduction of $37,851 (or 7.9%) from the Alperovits and Shamir solution of value $479,525. (The starting solution, feasible to problem NOP, had an objective value of $451,359, and was obtained by partially running the Alperovits and Shamir algorithm, itself started from the solution obtained by Fujiwara, et al. (1987). This reaffirms the importance of the possible synergy between models.) The run time was approximately 115 seconds.

If the original problem is modified to allow candidate pipe diameters having integer values from 1" to 24" (not necessarily all standard diameters), the solution given in Table 6.4 is obtained. This has a cost of $407,109 and yields a reduction of 9.6% from the revised solution of $450,249 obtained using the Alperovits and Shamir algorithm on the present problem that permits all the aforementioned pipe diameters. This solution also compares favorably with those obtained by Quindry, et al. (1979) and Fujiwara, et al. (1987), who report solutions having costs of $441,522 and $415,271, respectively. Overall, our solution yields a savings of $8,162 (2.0%) from the previously best known solution.

Table 6.1 Test Problem Node Data

Node	Elevation, m	Minimum Pressure Allowed	Supply or Demand
1	210	0	1120
2	150	30	-100
3	160	30	-100
4	155	30	-120
5	150	30	-270
6	165	30	-330
7	160	30	-200

Table 6.2 Test Problem Arc Data

Section	Arcs	Length, m	C_{HW}	Range of Allowable Diameters, in.	Selected Diameters, in.
1	(1,2)	1000	130	0-24	12, 14, 16, 18, 20
2	(2,3)	1000	130	0-24	6, 8, 10, 12, 14
3	(2,4)	1000	130	0-24	10, 12, 14, 16, 18
4	(4,5),(5,4)	1000	130	0-24	3, 4, 6, 8
5	(4,6),(6,4)	1000	130	0-24	10, 12, 14, 16, 18
6	(6,7),(7,6)	1000	130	0-24	8, 10, 12, 14, 16
7	(3,5),(5,3)	1000	130	0-24	6, 8, 10, 12, 14
8	(5,6),(7,5)	1000	130	0-24	6, 8, 10, 12, 14

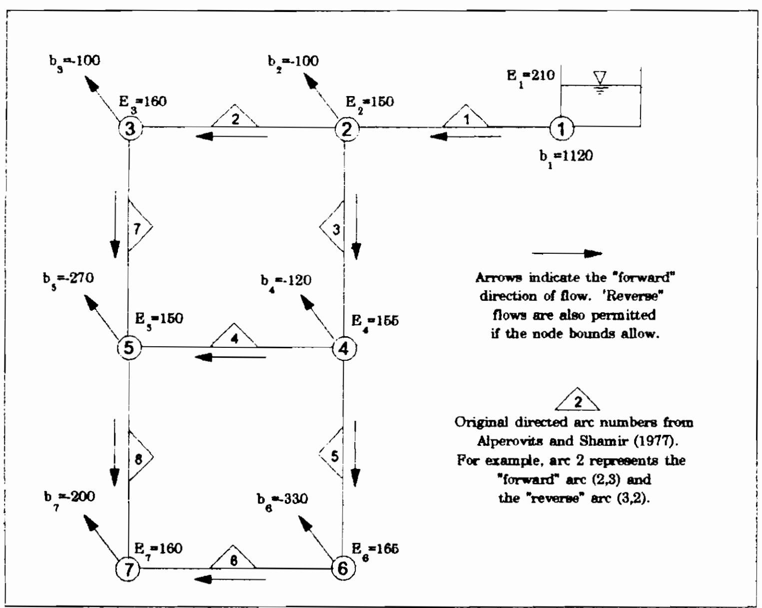

Figure 6.1 Test Problem Network Configuration

Table 6.3 Decomposition Method Solution for the Original Test Problem

Pipe Section # (i,j)	Sections having Length (m) of Diameter (in)	Flow (m^3/hr)	Head Loss (m)	Heads and Bounds			
				Node (i)	H_{iL} (m)	H_i+E_i (m)	H_{iU} (m)
1 (1,2)	569.0 18" 431.0 20"	1120.0	5.6	1	n/a	210.0	n/a
2 (2,3)	20.7 8" 979.3 10"	303.8	11.0	2	175.0	204.4	210.0
3 (2,4)	67.5 14" 932.5 16"	716.2	5.6	3	190.0	193.4	210.0
4 (4,5)	762.3 4" 237.7 6"	40.8	17.6	4	185.0	198.9	210.0
5 (4,6)	67.5 14" 932.5 16"	555.4	3.5	5	180.0	181.2	210.0
6 (6,7)	669.8 10" 330.2 12"	225.4	4.9	6	195.0	195.4	210.0
7 (3,5)	717.5 8" 282.5 10"	203.8	12.1	7	190.0	190.4	210.0
8 (7,5)	1000.0 4"	25.4	9.2				

Table 6.4 Decomposition Method Solution for the Modified Test Problem

Pipe Section # (i,j)	Sections having Length (m) of Diameter (in)	Flow (m^3/hr)	Head Loss (m)	Heads and Bounds			
				Node (i)	H_{iL} (m)	H_i+E_i (m)	H_{iU} (m)
1 (1,2)	235.9 18" 764.1 19"	1120.0	5.6	1	n/a	210.0	n/a
2 (2,3)	446.7 9" 553.3 10"	310.0	14.2	2	175.0	204.4	210.0
3 (2,4)	283.3 15" 716.7 16"	709.8	5.7	3	190.0	190.2	210.0
4 (4,5)	134.3 4" 865.7 5"	59.2	18.9	4	185.0	198.8	210.0
5 (4,6)	710.9 15" 289.1 16"	530.6	3.8	5	180.0	179.9	210.0
6 (6,7)	36.2 9" 963.8 10"	200.6	5.0	6	195.0	195.0	210.0
7 (3,5)	210.4 8" 789.6 9"	210.1	10.3	7	190.0	189.9	210.0
8 (7,5)	1000.0 1"	0.7	10.0				

6.2 Results from the RLT Global Optimization Algorithm of Chapter 5

The RLT strategy presented in Chapter 5 for solving the network optimization problem was also applied to the same test problem illustrated in Figure 6.1. This algorithm was implemented on a SUN SPARC 10 Unix workstation, using the CPLEX callable library to solve the linear programming subproblems. The RLT computer code is written in SUN FORTRAN 77, while the CPLEX code provided by the CPLEX Corporation is written in C.

Four different cases were tested using this algorithm, as applied to the original Alperovits and Shamir's (1977) test problem. Each case in this section represents a different set of initial bounds that are imposed on the flow variables. The four cases are illustrated in Figure 6.2 below, with the dark sections indicating the permissible ranges. Case A allows the maximum range of flows that appear to be viable for the problem on an initial inspection (no loop flows allowed, etc.). In this case, the flow is fixed at 1120 m^3/hr for arc 1, and the flow directions are determined for arcs 2 and 3. However, the directions for arcs 4-8 are not as yet determined, as represented by the dark rectangles extending on

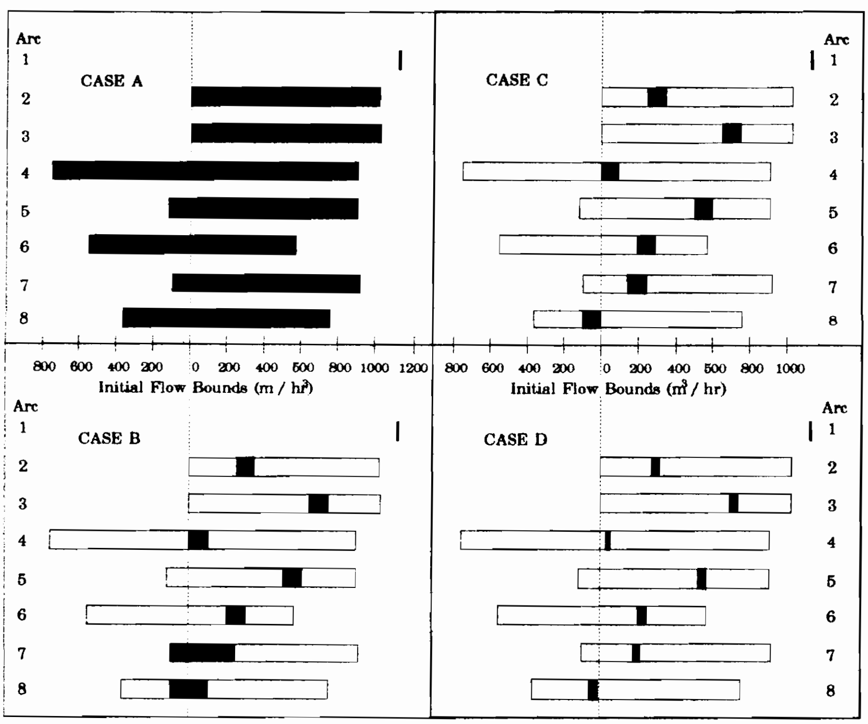

Figure 6.2 Initial Flow Bounds for Four Cases

both sides of the dashed line in the diagram. Case B restricts the flows greatly but still has two arcs (7 and 8) with flow directions as yet undetermined. Case C further restricts the variations in the flow and resolves the direction of all arc flows. Finally, Case D uses a very small hyperrectangle to restrict the flows, that is constructed around the favorable solution found in Section 6.1 above.

The run time dynamics for Case A are depicted below in Figure 6.3. The first linear programming relaxation for the entire feasible hyperrectangle has a value of \$226,418, which establishes a global lower bound for Problem NOP. As nodes are partitioned into smaller and smaller hyperrectangles, the least lower bound starts to increase as more fidelity is produced for the nonlinear approximations. Also, at each node of the branch-and-bound tree, we apply the following simple heuristic to possibly discover improved feasible solutions based on the lower bounding LP solutions. We fix the sizes of all the pipes at some percentage larger than those obtained by the LP relaxation by shifting pipe length allocations from smaller diameters to larger ones, and then solve the nonlinear system of equations that result upon fixing the x variables using the

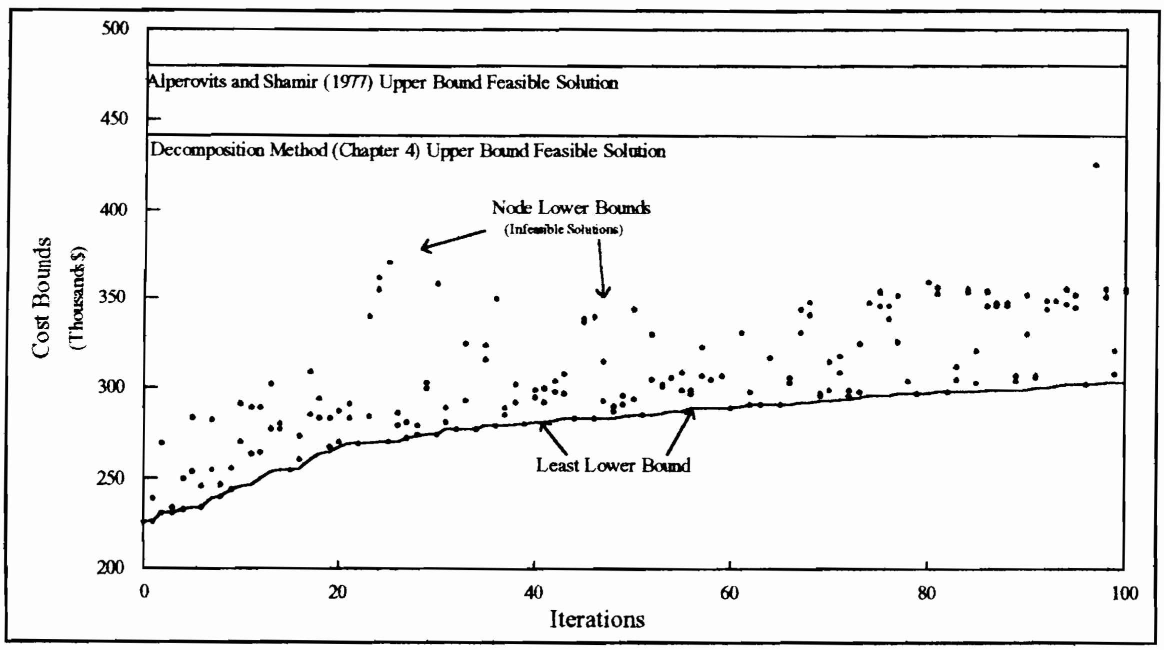

Figure 6.3 RLT Lower and Upper Bounds for Case A

linearization method recommended by Walski (1984). As the algorithm progresses, this percentage of increase in size is reduced as the problem becomes more restricted, in order to obtain better and better feasible upper bounding solutions. In this case, no feasible solutions were found in 100 iterations (35 minutes) by the heuristic applied at each node, and so, at this stage of the algorithm run, we have a global lower bound of $303,292 on the problem, and the present best known feasible solution (upper bound) has an objective value of $441,674. However, we note that we have been using a crude heuristic, our lower bounding problem can be further tightened as noted in Chapter 5, and that the solution of the bounding problem can be greatly accelerated by using Lagrangian relaxation techniques, instead of relying on CPLEX. (See related computations in Adams and Sherali (1993) and Sherali and Tuncbilek (1992, 1994) on using such techniques for some other classes of problems.)

For Case B, however, several feasible solutions were generated via the heuristic applied to the LP solution obtained at each branch-and-bound of each node, and so, we were able to further improve the upper bound on the problem (see Figure 6.4). In this case the RLT branch-and-bound methodology terminated with lower and upper bounds matching in 61 iterations (4 minutes), and hence optimally solved Problem NOP as restricted to the

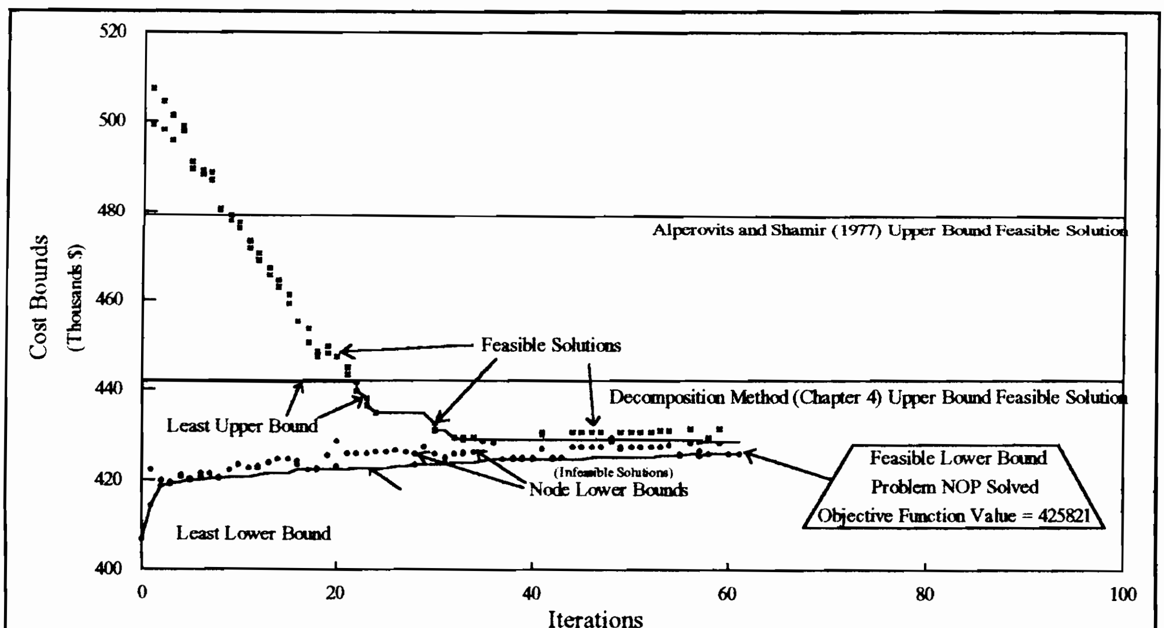

Figure 6.4 RLT Lower and Upper Bounds for Case B

specified hyperrectangle given in Figure 6.2. The corresponding optimal solution is presented in Table 6.5, and has a cost of $425,821. This improves upon the best known solution from Section 6.1 by another $15,853 (3.5%).

Table 6.5 RLT Method Solution for the Test Problem (Case B)

Pipe Section # (i,j)	Sections having Length (m) of Diameter (in)	Flow (m^3/hr)	Head Loss (m)	Heads and Bounds Node (i)	H_{iL} (m)	H_i+E_i (m)	H_{iU} (m)
1 (1,2)	984.0 18" 16.0 20"	1120.0	6.7	1	n/a	210.0	n/a
2 (2,3)	1000.0 10"	320.0	11.6	2	175.0	203.3	210.0
3 (2,4)	1000.0 16"	700.0	5.0	3	190.0	191.7	210.0
4 (4,5)	445.0 3" 555.0 4"	23.0	18.1	4	185.0	198.3	210.0
5 (4,6)	1000.0 16"	557.0	3.3	5	180.0	180.0	210.0
6 (6,7)	819.0 10" 181.0 12"	227.0	5.5	6	195.0	195.0	210.0
7 (3,5)	520.0 8" 480.0 10"	220.0	11.7	7	190.0	190.0	210.0
8 (7,5)	1000.0 4"	27.0	10.3				

Figure 6.5 illustrates the run time dynamics for Case C. As the initial hyperrectangle was more restrictive, we notice that the initial lower bound as well as the final solution value obtained are both higher than those for Case B. However, the run time is longer, which was not expected. We have verified that the reason is that there exist many alternative optimal solutions to the LP relaxations, and so the overall effort is strongly influenced by the branching strategy, and how quickly this is able to resolve such alternative relaxed solutions that do not correspond to improving feasible designs.

For Case D, we tried an interesting experiment. Running this case as usual, the problem found an optimal solution that a slightly better objective value, after performing 33 iterations. However, we next tightened the relaxation by introducing the additional RLT constraints of the type (5.1c)*H_i as recommended in Sections 5.3 and 5.8. Once these constraints were added to the formulation, the resulting initial LP relaxation turned out to have an optimal corner point at a feasible solution. The nonconvex problem was therefore solved at the first node itself, via a single linear program! This reinforces our contention that the proposed RLT method can greatly benefit from the recommended strategies for tightening the LP relaxations.

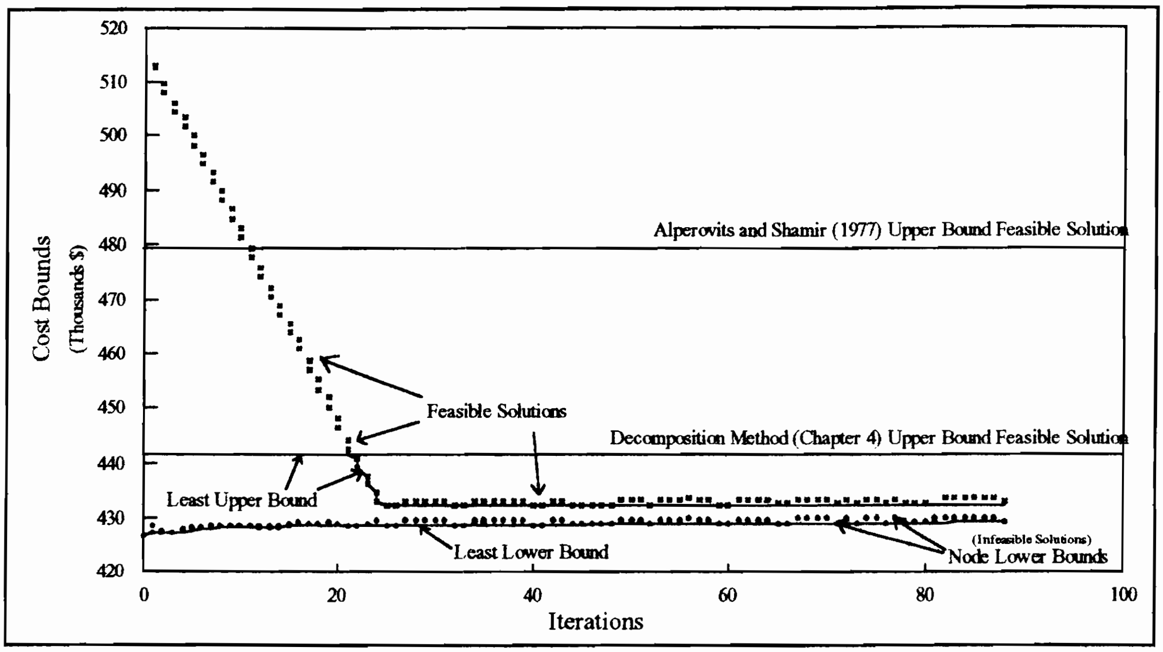

Figure 6.5 RLT Lower and Upper Bounds for Case C

6.3 RLT Results for the Design Problem Having Integer Exponents in the Hazen-Williams Equations

The results of applying the RLT branch-and-bound algorithm on the same test problem, but using the integer exponent version of the Hazen-Williams equation (3.3c), are presented in Figure 6.6. The initial flow restrictions for Case B in section 6.2 were used for the initial hyperrectangle. The solution presented in Table 6.6 is within 5% of global optimality and was obtained in 87 iterations of the algorithm. Note that its cost of \$504,046 is significantly larger than the comparable one found in Table 6.5 above for the exponent of 1.852. In fact, the integer exponent solutions will, in general, produce more expensive solutions since the exponent of 2 in the Hazen-Williams equation results in greater head losses than does the exponent of 1.852. The ensuing compensation between flow rates and increased pipe diameters in order to maintain adequate pressure heads, naturally produces a more expensive design using the larger (integer) exponent.

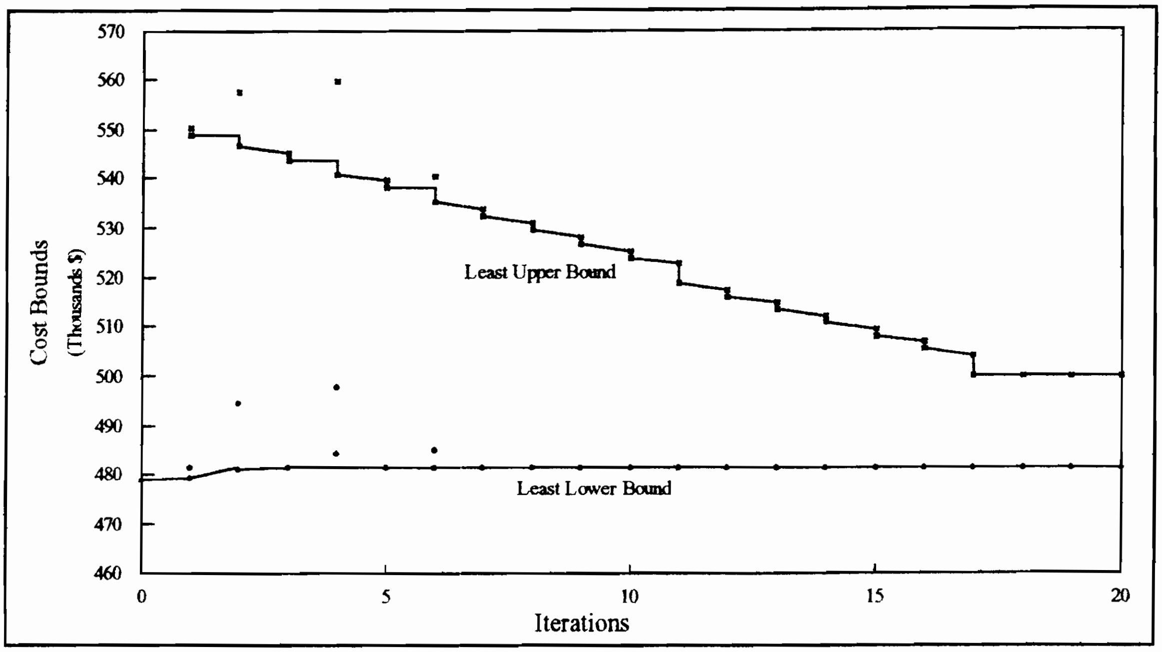

Figure 6.6 RLT Lower and Upper Bounds for the Test Problem with Integer Exponents

Table 6.6 RLT Method Solution for the Test Problem with Integer Exponents (Case B)

Pipe Section # (i,j)	Sections having Length (m) of Diameter (in)	Flow (m^3/hr)	Head Loss (m)	Heads and Bounds: Node (i)	H_{iL} (m)	H_i+E_i (m)	H_{iU} (m)
1 (1,2)	1000. 20"	1120.	5.8	1	n/a	210.	n/a
2 (2,3)	763. 10" 237. 12"	325.	12.2	2	175.	204.	210.
3 (2,4)	482. 16" 518. 18"	695.	5.1	3	190.	192.	210.
4 (4,5)	830. 3" 170. 4"	19.	14.9	4	185.	199.	210.
5 (4,6)	830. 16" 170. 18"	555.	8.6	5	180.	184.	210.
6 (6,7)	671. 10" 224. 12" 105. 14"	225.	5.2	6	195.	191.	210.
7 (3,5)	119. 8" 881. 10"	225.	8.4	7	190.	186.	210.
8 (7,5)	1000.0 4"	25.	2.6				

7. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

7.1 Summary and Significance of Results

In this research effort, we have proposed for the first time, an approach that integrates pipe reliability-and-cost and pipe network optimization models within a network design process. In this scheme, the results from each submodel are used in a feedback loop connected into the other model, until a stable design is attained. The pipe reliability and cost submodel uses statistical methodologies based on historical records of pipe breaks to estimate the reliabilities of individual pipe segments, and hence to estimate future annualized costs. The pipe network design submodel provides replacement and new construction decisions for a variety of demand patterns, while incorporating a level of hydraulic redundancy which ensures that these demands can be satisfied with any one pipe removed from the network. A network optimization routine is used to solve for each network configuration and demand pattern. For these hard nonconvex optimization problems, we have developed two new approaches that permit us to take advantage of the special network structures inherent within the problem. Namely, we have presented a variable projection/ decomposition approach that leads to better locally optimal solutions, and a new Reformulation-Linearization Technique that provides global optimal solutions, available for the first time for these types of problems. The integrated design approach constructs a hydraulically redundant network, resulting in a comprehensive reduced cost water distribution system that meets all pressure and flow requirements for realistic problems, even under a wide variety of pipe failure modes.

Preliminary results in implementing the proposed pipe network design submodel using both the decomposition and RLT strategies have shown encouraging improvements in solutions obtained for a simple single demand pattern test problem that has been attempted by several researchers. In particular, for this example, our approach detects several solutions that significantly improve over the best known solution available in the literature. Furthermore, we have been able to provide a capability for determining valid global lower bounds for the overall problem, and in particular, to globally solve certain instances of restricted flow problems.

We have also found that certain additional RLT constraint sets that are specially tailored to the network flow structure in the problem can significantly improve the performance of the RLT methodology by providing for tighter relaxations. Even though these constraints are not required for convergence, they can reduce certain branch-and-bound search subtrees containing many nodes down to a single node.

The advantages of our formulation and integrated design approach are as follows:

(1) Synergy between pipe-reliability and cost models, and network design models.
(2) The life cycle costs of all pipes are taken into consideration (annualized capital and maintenance costs) allowing the comparison between retaining existing pipes and replacing them with new ones.
(3) Existing pipe links may remain intact or may be either partially or completely replaced.
(4) Multiple pipe segments of fixed diameter, roughness, and cost are admissible for composing each network link.
(5) The network layout is not limited to tree graphs at any stage in the design/optimization process. Loops are allowed during all stages.
(6) Pipe segments are properly limited to commercially available diameters, and any combination of adjacent diameters is allowed for each link. That is, links are not artificially limited to a small selection of available diameters, as in existing practice, in order to reduce the number of variables in the problem. However, as seen in the test problem, this *capability* is also available if desired.
(7) Links can be removed from the layout at any particular stage in the design/optimization process if it is advantageous to do so.
(8) No *a priori* knowledge of any flows, pipe sizes (other than existing pipes), or directions of flow is required.
(9) Energy source head levels are determined in the design process based on the increased cost to raise the head from a baseline level or elevation.
(10) Individual pipe segment cost and friction coefficients are incorporated within the analysis.
(11) Estimates of individual pipe segment reliabilities are incorporated.
(12) Hydraulic pressure and flow feasibility are retained for all demand patterns and with any one link removed from the system for failure or maintenance reasons. Hence, real

hydraulic network reliability is addressed in a concrete physical fashion without neglecting the optimization process.

(13) Advanced decomposition methods are used to project the optimization problems at each stage onto a series of reduced subspaces where the subproblems are easy to solve. This improves on the solutions obtained and may enhance large-scale problem solvability.

(14) Global optimization schemes are considered for individual stage problems, that will ensure the merit of solutions obtained. No competing procedure that can guarantee global optimality exists in the literature.

We believe, in fact, that items (1), (3), (7), (8), (12) - (14) are being incorporated into working algorithms for the first time, each having a significant impact on the ability to solve "real world" problems. We believe that since so many of these capabilities are included, that new more realistic test problems should be developed that exercise these capabilities, and hence, be able to even better discriminate between existing methods and the newer methods presented here.

7.2 Conclusions and Recommendations for Future Research

Our recommendations for future research lie primarily in two areas. First, there are several tasks left undone at this time in the implementation of the integrated design approach as presented in Chapter 3. The most important of these is the design for redundancy which is accomplished by iteratively redesigning the network with each arc removed, and incrementally updating the overall design after each iteration. Multiple moves that permit one to "jump" from one network structure to a modified architecture (with links removed) where lower cost feasible solutions might exist, are also worth investigating. We believe that this process will make strong gains toward reaching the goal of a realistic design in the spirit of the discussion in Chapter 1, while keeping optimization methods (versus ad hoc heuristics) as the underlying drivers in the network design process.

Further research can also take place in the packaging and automation of the integrated design process in a user-friendly interactive manner, including the consideration of alternate parameter choices as outlined in Section 3.4. These parameters help to custom-fit the design process to the local operating procedures and preferences of the decision maker. Such a capability can provide considerable insight for practitioners.

The second area for future research lies in the modification of the RLT global optimization algorithm. There are five major improvements that we predict will significantly improve its performance. First, is the inclusion of tightening constraints such as those previously mentioned of the type $(5.1c)^*H_i$. As noted before, these constraints made notable gains in the speed of convergence for certain constrained problems and will no doubt improve the overall performance for other such problems as well. More generally, constraint sets that take advantage of the inherent problem structure could be sought out to further tighten the LP relaxations, so that fewer nodes need be processed in the branch-and-bound tree. In the same spirit, classes of constraints that are being currently generated but do not significantly contribute to the tightening of the relaxation to the theoretical convergence of the algorithm, can be identified and deleted from the formulation.

Second, better heuristics need to be developed (or implemented) that can quickly find good quality feasible solutions when starting from an infeasible solution obtained via a lower bounding problem solved at a node in the branch-and-bound tree. We only just began to investigate this area, and the heuristic we implemented was a crude preliminary attempt. However, it is encouraging that this same simple heuristic was able to generate feasible solutions of good enough quality to quickly reduce the lower-upper bound gap to within 0-5% in certain cases. If more powerful heuristics were developed and integrated with stronger lower bounding techniques as above, then this might provide the capability for solving realistic problems to at least within 5% of optimality with reasonable effort.

Third, there exist several alternative admissible branching strategies that can lead to theoretical convergence. Such strategies should be identified and computationally tested.

Fourth, there exist several "postprocessing" steps that can tighten the bounding hyperrectangle of flows after the LP relaxation is solved at any node of the branch and bound tree. The idea is as follows. Suppose that we know an incumbent solution to problem NOP having an objective value of v^*. Restricting the objective function $c{\cdot}x + C{\cdot}X + c_s{\cdot}H_s$ to a value less than or equal to v^* *as a constraint*, we change the objective function to alternately maximizing one of the flow variables Q_{ij} for one LP, and then minimize it for another. This would yield a tighter range of values for this Q_{ij} variable. Reconstructing the RLT formulation using these revised bounds, we repeat the process for each flow variable Q_{ij} in turn, each time potentially tightening the formulation.

Once all the flow bounds have been thus tightened, we can cyclically repeat the process for a limited number of times, or until no more than a, say, 10% reduction in the volume of the bounding hyperrectangle is achieved, whichever occurs first. This will certainly improve the performance of the RLT branch-and-bound procedure even if only applied at the initial node. Perhaps this process could also be repeated early in the algorithm while the heuristic for generating feasible solutions is suppressed, until the lower-upper bound gap is reduced to some level, say 10%. At such a time, the extra work could be transferred from this range reduction strategy that could be turned off, to the heuristic feasible solution generator which can then be switched on. In this manner, we might expect improved quality solutions to be produced by the heuristic procedure.

Finally, we could employ nonlinear outer approximations to the feasible region that generate nonlinear, but convex, branch-and-bound node subproblems. Thus the nonconvex problem NOP could be reformulated and then transformed into a series of *nonlinear* convex subproblems (perhaps called the Reformulation-Convexification Technique (RCT)) instead of a series of *linear* convex subproblems as we have employed. In such a case, perhaps a method could be found that is guaranteed to theoretically converge to a globally optimal solution within finite time. Computer implementations of such an algorithm would be unquestionably faster than the RLT method employed herein.

BIBLIOGRAPHY

Adams, W. P., and Johnson, T. A., "Improved Linear Programming-Based Lower Bounds for the Quadratic Assignment Problem," Manuscript, 1994.

Adams, R. W., "Distribution Analysis by Electronic Computer," Institution of Water Engineers, Vol. 15, 1961, pp. 415-428.

Adams, W. P., and Sherali, H. D., "A Tight Linearization and an Algorithm for Zero-One Quadratic Programming Problems," Management Science, Vol. 32, No. 10., 1986, pp. 1274-1290.

Adams, W. P., and Sherali, H. D., "Mixed-Integer Bilinear Programming Problems," Mathematical Programming, Vol. 59, No. 3, 1993, pp. 279-305.

Alperovits, E., and Shamir, U., "Design of Optimal Water Distribution Systems," Water Resources Research, Vol. 13, December 1977, pp. 885-900.

Andreou, S. A., Marks, D. H., and Clark, R. M., "A New Methodology for Modelling Break Failure Patterns in Deteriorating Water Distribution Systems: Theory," Advances in Water Resources, Vol. 10, March 1987a, pp. 2-10.

Andreou, S. A., Marks, D. H., and Clark, R. M., "A New Methodology for Modelling Break Failure Patterns in Deteriorating Water Distribution Systems: Applications," Advances in Water Resources, Vol. 10, March 1987b, pp. 11-21.

Bazaraa, M. S., Jarvis, J. J., and Sherali, H. D., Linear Programming and Network Flows, second edition, John Wiley and Sons, New York, 1990.

Bazaraa, M. S., Sherali, H. D., and Shetty, C. M., Nonlinear Programming: Theory and Algorithms, second edition, John Wiley and Sons, New York, 1993.

Bhave, P. R., "Noncomputer Optimization of Single-Source Networks," Journal of the Environmental Engineering Division, ASCE, Vol. 104, No. EE4, August 1978, pp. 799-813.

Bhave, P. R., "Optimization of Gravity-Fed Water Distribution Systems: Theory," Journal of Environmental Engineering, ASCE, Vol. 109, No. 1, February 1983a, pp. 189-205.

Bhave, P. R., "Optimization of Gravity-Fed Water Distribution Systems: Application," Journal of Environmental Engineering, ASCE, Vol. 109, No. 2, April 1983b, pp. 383-395.

Bhave, P. R., "Optimal Expansion of Water Distribution Systems," Journal of the Environmental Engineering Division, ASCE, Vol. 111, No. EE2, 1985, pp. 177-197.

Clark, R. M., Stafford, C. L., and Goodrich, J. A., "Water Distribution Systems: A Spatial and Cost Evaluation," Journal of the Water Resources Planning and Management Division, ASCE, Vol. 1008, No. WR3, October 1982, pp. 243-256.

Collins, M., Cooper, L., Helgason, R., Kennington, J., and LeBlanc, L., "Solving the Pipe Network Analysis Problem Using Optimization Techniques," Management Science, Vol. 24, No. 7, March 1978, pp. 747-760.

Cross, H., "Analysis of Flow in Networks of Conduits or Conductors," Engineering Experiment Station Bulletin No. 286, University of Illinois, 1936.

Delfino, W. C. D., "Optimal Design of Water Distribution Pipeline Networks," Ph.D. thesis, Case Western Reserve University, June 1975.

Eggener, C. L., and Polkowski, L. B., "Network Models and the Impact of Model Assumptions," Journal of the American Water Works Association, Vol. 68, 1976, pp. 189-196.

Eiger, G., Shamir, U., and Ben-Tal, A., "Optimal Design of Water Distribution Networks," Water Resources Research, Vol. 30, No. 9, 1994, pp. 2637-2646.

Fair, G., M., Geyer, J. C., and Okum, D. A., Water and Wastewater Engineering, Vol. 1, Water Supply and Wastewater Removal, Wiley, New York, 1966.

Fiacco, A. V., and McCormick, G. P., "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal-Dual Method," Management Science, Vol. 10, 1964a, pp. 360-366.

Fiacco, A. V., and McCormick, G. P., "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming," Management Science, Vol. 10, 1964b, pp. 601-617.

Fujiwara, O., and Dey, D., "Two Adjacent Pipe Diameters at the Optimal Solution in the Water Distribution Network Models," Water Resources Research, Vol. 23, No. 8, August 1987, pp. 1457-1460.

Fujiwara, O., Jenchaimahakoon, B., and Edirisinghe, N. C. P., "A Modified Linear Programming Gradient Method for Optimal Design of Looped Water Distribution Networks," Water Resources Research, Vol. 23, No. 6, June 1987, pp. 977-982.

Fujiwara, O., and Khang, D. B., "A Two-Phase Decomposition Method for Optimal Design of Looped Water Distribution Networks," Water Resources Research, Vol. 26, No. 4, April 1990, pp. 539-549.

Fujiwara, O., and Tung, H. D., "Reliability Improvement for Water Distribution Networks Through Increasing Pipe Size," Water Resources Research, Vol. 27, No. 7, July 1991, pp. 1395-1402.

Gessler, J. M., "Optimization of Pipe Networks," paper presented at Proceedings of the Ninth International Symposium on Urban Hydrology, Hydraulics and Sediment Control, University of Kentucky,Lexington, July 27-30, 1982.

Gessler, J. M., "Pipe Network Optimization by Enumeration," Computer Applications in Water Resources, ed. H. C. Torno, 1985, pp. 572-581.

Glover, F., and Wolsey, E., "Converting the 0-1 Polynomial Programming Problem to a 0-1 Linear Program," Operations Research, Vol. 22, No. 1, 1974, pp. 180-182.

Glover, F., "Improved Linear Integer Programming Formulations of Nonlinear Integer Problems," Management Science, Vol. 22, No. 4, December 1975, pp. 455-460.

Grant, E. L., Ireson, W. G., and Leavenworth, R. S., Principles of Engineering Economy, seventh edition, John Wiley and Sons, New York, 1982.

Hestenes, M. R., and Stiefel, E., "Methods of Conjugate Gradients for Solving Linear Systems," Journal Res. National Bureau Standards Section B, Vol. 48, 1952, pp. 409-436.

Hobbs, B. F., and Hepenstal, A., "Is Optimization Optimistically Biased?," Water Resources Research, Vol. 25, No. 2, February 1989, pp. 152-160.

Hoffman, K. L., and Padberg, M., "Improving LP-Representations of Zero-One Linear Programs for Branch-and_Cut," ORSA Journal on Computing, Vol. 3, No. 2, 1991, pp. 121-134.

Horst, R. and Tuy, H., Global Optimization: Deterministic Approaches, Springer Verlag, Berlin, second edition, 1993.

Jacoby, S. L. S., "Design of Optimal Hydraulic Networks," Journal of Hydraulics Division, ASCE, Vol. 94, HY3, 1968, pp. 641-666.

Jeppson, R. W., "Practical Optimization of Looped Water Systems," in Computer Applications in Water Resources, ed. H. C. Tonro, 1985, pp. 723-731.

Kapur, K. C., and Lamberson, L. R., Reliability in Engineering Design, John Wiley and Sons, New York, 1977.

Karaa, F. A., Marks, D. H., and Clark, R. M., "Budgeting of Water Distribution Improvement Projects," Journal for Water Resources Planning and Management, ASCE, Vol. 113, No. 3, May 1987, pp. 378-391.

Kessler, A., and Shamir, U., "Analysis of the Linear Programming Gradient Method for Optimal Design Of Water Supply Networks," Water Resources Research, Vol. 27, No. 7, July 1989, pp. 1469-1480.

Kettler, A. J., and Goulter, I. C., "An Analysis of Pipe Breakage in Urban Water Distribution Networks," Canadian Journal of Civil Engineering, Vol. 12, 1985, pp. 286-293.

Lansey, K., and Mays, L., "A Methodology for Optimal Network Design," Computer Applications in Water Resources, ed. H. C. Torno, 1985, pp. 732-738.

Li, D., and Haimes, Y. Y., "Optimal Maintenance-Related Decision Making for Deteriorating Water Distribution Systems, 1. Semi-Markovian Model for a Water Main," Water Resources Research, Vol. 28, No. 4, April 1992a, pp. 1053-1061.

Li, D., and Haimes, Y. Y., "Optimal Maintenance-Related Decision Making for Deteriorating Water Distribution Systems, 2. Multilevel Decomposition Approach," Water Resources Research, Vol. 28, No. 4, April 1992b, pp. 1063-1070.

Loganathan, G. V., Sherali, H. D., and Shah, M. P.m "A Two-Phase Network Design Heuristic for Minimum Cost Water Distribution System Under a Reliability Constraint," Engineering Optimization, Vol. 15(4), 1990, pp. 311-336.

Loubser, B. F., and Gessler, J. M., "Computer Aided Optimization of Water Distribution Networks," Integrated Computer Applications in Water Supply, Vol. 1, Research Studies Press Ltd., Somerset, England, 1993, pp. 103-120.

Mays, L. W., Reliability Analysis of Water Distribution Systems, American Society of Civil Engineers, New York, NY, 1989.

Morgan, D. R., and Goulter, I. C., "Optimal Urban Water Distribution Design," Water Resources Research, Vol. 21, No. 5, May 1985, pp. 642-652.

Nemhauser, G. L., and Wolsey, L. A., Integer and Combinatorial Optimization, John Wiley and Sons, New York, NY, 1988.

Quindry, G., Brill, E. D., and Liebman, J. C., "Optimization of Looped Water Distribution Systems," Journal of Environmental Engineering Division, ASCE, Vol. 107, EE4, August 1981, pp. 665-679.

Quindry, G., Brill, E. D., Liebman, J. C., and Robinson, A., "Comments on Design of Optimal Water Distribution Systems," by E. Alperovits and U. Shamir, Water Resources Research, Vol. 15, No. 6, December 1979, pp. 1651-1654.

Rothfarb, B., Frank, H., Rosenbaum, D. M., Steiglitz, K., and Kleitman, D. J., "Optimal Design of Offshore Natural-Gas Pipeline Systems," Operations Research, Vol. 18, November-December 1970, pp. 992-1020.

Rowell, W. F., "A Methodology for Optimal Design of Water Distribution Systems," Ph.D. thesis, University of Texas at Austin, 1979.

Rowell, W. F., and Barnes, J. W., "Obtaining Layout of Water Distribution Systems," Journal of the Hydraulics Division, ASCE, Vol. 108, HY1, 1982, pp. 137-148.

Sellers, C., D., "Discussion of Obtaining the Layout of Water Distribution Systems," by W. F. Rowell and J. W. Barnes, Journal of the Hydraulics Division, ASCE, Vol. 110, HY1, 1984, pp. 68-69.

Shaake, J. C., Jr., and Lai, D., "Linear Programming and Dynamic Programming Application of Water Distribution Network Design," Rep. 116, Hydrodynamic Lab., Department of Civil Engineering, MIT Press, Cambridge, Mass., 1969.

Shamir, U., and Howard, C. D. D., "An Analytic Approach to Scheduling Pipe Replacement," Journal of the American Water Works Association, May 1979, pp. 248-258.

Shamir, U., and Howard, C. D. D., "Water Supply Reliability Theory," Journal of the American Water Works Association, July 1981, pp. 379-384.

Sherali, H. D., "Tight Linear Programming Representations: A Key to Solving Nonconvex Programming Problems," Invited Colloquium Paper in Challenges for Computational Mathematics in the 1990's, eds. M. J. Saltzman and D. Warner, to appear, Manuscript, 1991.

Sherali, H. D., and Adams, W. P., "A Decomposition Algorithm for a Discrete Location-Allocation Problem," Operations Research, Vol. 32, No. 4, 1984, pp. 878-900.

Sherali, H. D., and Adams, W. P., "A Hierarchy of Relaxations and Convex Hull Representations for Mixed-Integer Zero-One Programming Problems," Discrete Applied Mathematics, Vol. 52, 1994, pp. 83-106.

Sherali, H. D., and Adams, W. P., "A Hierarchy of Relaxations Between the Continuous and Convex Hull Representations for Zero-One Programming Problems," SIAM Journal on Discrete Mathematics, Vol. 3, No. 3, pp. 411-430, 1990.

Sherali, H. D., and Adams, W. P., "A Reformulation-Linearization Technique (RLT) for Solving Discrete and Continuous Nonconvex Programming Problems," Mathematics Today, to appear, Manuscript, 1993.

Sherali, H. D., and Alameddine, A., "A New Reformulation-Linearization Technique for Solving Bilinear Programming Problems," Journal of Global Optimization, Vol. 2, 1992, pp. 379-410.

Sherali, H. D., and Myers, D. C., "Dual Formulations and Subgradient Optimization Strategies for Linear Programming Relaxations of Mixed-Integer Programs," Discrete Applied Mathematics, 20(S-16), 1989, pp. 51-68.

Sherali, H. D., Ramachandran, S., and Kim, S., "A Localization and Reformulation Discrete Programming Approach for the Rectilinear Distance Location-Allocation Problem," Discrete Applied Mathematics, Vol. 49, Nos 1-3, 1994, pp. 357-378.

Sherali, H. D., and Smith, E. P., "An Optimal Replacement-Design Model for a Reliable Water Distribution Network System," Integrated Computer Applications in Water Supply, Vol. 1, Research Studies Press Ltd., Somerset, England, 1993, pp. 61-75.

Sherali, H. D., and Tuncbilek, C. H., "A Global Optimization Algorithm for Polynomial Programming Problems Using a Reformulation-Linearization Technique," Journal of Global Optimization, Vol. 2, 1992a, pp. 101-112.

Sherali, H. D., and Tuncbilek, C. H., "A Squared-Euclidean Distance Location-Allocation Problem," Naval Research Logistics, Vol. 39, 1992b, pp. 447-469.

Sherali, H. D., and Tuncbilek, C. H., "A Reformulation-Convexification Approach for Solving Nonconvex Quadratic Programming Problems," Manuscript, Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, 1994.

Sherali, H. D., and Ulular, O., "Conjugate Gradient Methods Using Quasi-Newton Updates with Inexact Line Searches," Journal of Mathematical Analysis and Applications, Vol. 150, No. 2, August 1990, pp. 359-377.

Stacha, J. H., "Criteria for Pipeline Replacement, "Journal of the American Water Works Association, May 1978, pp. 256-258.

Su, Y. C., Mays, L. W., Duan, N., and Lansey, K. E., "Reliability Based Optimization Model for Water Distribution Systems," Journal of Hydraulic Engineering, ASCE, Vol. 114, No. 12, 1987, pp. 1539-1556.

Sullivan, J. P., Jr., "Maintaining Aging Systems - Boston's Approach," Journal of the American Water Works Association, November 1982, pp. 555-559.

Templeman, B., "Discussion of Optimization of Looped Water Distribution Systems" by Quindry et al., Journal of the Environmental Engineering Division, ASCE, Vol. 108, EE3, 1982, pp. 599-602.

Wagner, J. M., Shamir, U., and Marks, D. H., "Water Distribution Reliability: Analytical Methods," Journal of Water Resources Planning and Management, ASCE, Vol. 114, No. 3, May 1988a, pp 253-275.

Wagner, J. M., Shamir, U., and Marks, D. H., "Water Distribution Reliability: Simulation Methods," Journal of Water Resources Planning and Management, ASCE, Vol. 114, No. 3, May 1988b, pp 276-294.

Walski, T. M., Analysis of Water Distribution Systems, Van Nostrand Reinhold Company, New York, 1984.

Walski, T. M., "State-of-the-Art Pipe Network Optimization," in Computer Applications in Water Resources, ed. H. C. Tonro, 1985, pp. 559-568.

Walski, T. M., Water Supply System Rehabilitation, American Society of Civil Engineers, New York, 1987.

Wood, D. J., and Charles, C. O., "Hydraulic Network Analysis Using Linear Network Theory," Journal of the Hydraulics Division, ASCE, Vol. 98, HY7, July 1972, pp. 1157-1171.

VITA

Ernest Price Smith was born in Alexandria, Virginia, on September 29, 1961, the son of Carey and Fan Smith. After completing his studies at McLean High School in McLean, Virginia, in 1979, he began studies at Virginia Polytechnic Institute and State University. During his third year at Virginia Tech, he enlisted in the United States Air Force as an Airman First Class in the College Senior Engineering Program. Subsequently, he married Laurie Bear of Sylva, North Carolina on August 21, 1982. After graduating from Virginia Tech in December, 1982 with a Bachelor of Science in Electrical Engineering, he attended Officer Training School in San Antonio, Texas, where he was commissioned as a second lieutenant in the United States Air Force on April 1, 1983. For the next three years he worked as an electronic warfare test engineer at Wright Patterson Air Force Base in Dayton, Ohio. He was awarded the degree of Master of Science in Systems Engineering from the Air Force Institute of Technology in December, 1987. From 1988 until 1991 he worked as a space and electronics test analyst at the Air Force Operational Test and Evaluation Center at Kirtland Air Force Base, Albuquerque, New Mexico. In August, 1991 he returned to Virginia Tech in the Department of Industrial and Systems Engineering for graduate school under the Air Force Civilian Institution Program. He has two children, Stephen Carey, age 7, and Rachel Lynn, age 3. His next military assignment will be at the Air Force Institute of Technology as an Assistant Professor of Operations Research in the Department of Operational Sciences.

Permanent Address: 9714 West Route 36
St. Paris, OH 43072

Ernest Price Smith