Safe and Tight Linear Estimators for Global Optimization

Glencora Borradaile and Pascal Van Hentenryck

Brown University
Box 1910
Providence, RI, 02912
{glencora, pvh}@cs.brown.edu

Abstract. Global optimization problems are often approached by branch and bound algorithms which use linear relaxations of the nonlinear constraints computed from the current variable bounds. This paper studies how to derive safe linear relaxations to account for numerical errors arising when computing the linear coefficients. It first proposes two classes of safe linear estimators for univariate functions. Class-1 estimators generalize previously suggested estimators from quadratic to arbitrary functions, while class-2 estimators are novel. Class-2 estimators are shown to be tighter theoretically (in a certain sense) and almost always tighter numerically. The paper then generalizes these results to multivariate functions. It shows how to derive estimators for multivariate functions by combining univariate estimators derived for each variable independently. The combination of tight class-1 safe univariate estimators is shown to be a tight class-1 safe multivariate estimator and multivariate class-2 estimators are shown to be theoretically tighter (in a certain sense) than multivariate class-1 estimators. Finally, the paper describes how to apply these estimators to approximate linear programs with interval coefficients safely.

1 Introduction

Global optimization problems arise naturally in many application areas, including chemical and electrical engineering, biology, economics, and robotics to name only a few. They consist of finding all solutions or the global optima to nonlinear programming problems. These problems are inherently difficult computationally (i.e., they are PSPACE-hard) and may also be challenging numerically. In addition, there has been considerable interest in recent years to produce rigorous or reliable results, i.e., to make sure that the exact solutions are enclosed in the results of the algorithms.

Global optimization problems are often approached by branch and bound algorithms which use linear relaxations of the nonlinear constraints computed from the current bounds for the variables at each node of the search tree (e.g., [3, 5, 6, 11, 12, 15, 16]). The linear relaxation can be used to obtain a lower bound on the objective function (in minimization problems) and/or to update the variable bounds. This approach can also be combined with constraint satisfaction techniques for global optimization which are also effective in reducing the variable bounds (e.g., [1, 2, 7, 13, 14]).

The linear relaxation is generally obtained by linearizing nonlinear terms independently, giving what is often called linear over- and under-estimators. When rigorous and reliable results are desired, it is critical to generate a safe linear relaxation which over-approximates the solution to the nonlinear problem at hand (e.g., [9, 10]). Indeed, the coefficients in the linear constraints are generally given by real functions which are subject to rounding errors when evaluated. As a consequence, the resulting linear relaxation may not be safe. Moreover, naive approaches (e.g., upward rounding for the overestimators' coefficients) are not safe in general either. Once a safe linear relaxation is available, it can be solved exactly or safe bounds on the objective function can be obtained using duality as in [10, 4] for instance. Experimental results (e.g., [9]) have shown that both of these two corrections are critical in practice, even on simple problems, to find all solutions to nonlinear polynomial systems.

This paper focuses on obtaining safe linear relaxations for global optimization problems and contains three main contributions:

1. The paper presents two classes of safe estimators of univariate functions. The first class of estimators generalizes the results of [9] from quadratic to arbitrary functions, while the second class is entirely novel. Theoretical tightness results are given for both classes, giving the relative strengths of the presented estimators. In particular, the results show that class-2 estimators (when they apply) are theoretically tighter than class-1 estimators (in a certain sense to be defined). Moreover, the numerical results indicate that class-2 estimators are almost always tighter in practice in our experiments.
2. The paper then generalizes the univariate results to multivariate functions. It shows how to derive estimators for multivariate functions by combining univariate estimators derived for each variable independently. Moreover, the combination of tight class-1 safe univariate estimators is shown to give a tight class-1 safe multivariate estimator. Finally, univariate relative tightness results are shown to carry over to the multivariate case, i.e., multivariate class-2 estimators are shown theoretically tighter (in a certain sense) than multivariate class-1 estimators.
3. The paper also shows how to approximate linear programs with interval coefficient safely. This is also important to guarantee correctness, since coefficients are generally not known with certainty or are given by expressions or textual representations which are subject to rounding errors when evaluated. Interestingly, linear programs with interval coefficients can be approximated by class-1 and class-2 estimators again.

As a consequence, these results provide a systematic, comprehensive, and elegant framework to derive safe linear estimators for global optimization problems. In conjunction with the safe bounds on the linear relaxations derived in [10, 4], they provide the theoretical foundation for rigorous results in branch and bound approaches to global optimization based on linear programming.

The rest of this paper is organized as follows: Section 2 defines the concept of safe estimators and the problems arising in deriving them. Section 3 derives the two classes of linear estimators for univariate functions. Section 4 presents the theoretical and numerical tightness results. Section 5 presents the multivariate results. Section 6 shows how to safely approximate linear programs with interval coefficients.

2 Definitions and Problem Statement

This section defines the concepts of linear estimators and safe linear overestimators, as well as the problem tackled in this paper. For simplicity, the definitions are given for univariate functions only, but they generalize naturally to multivariate functions. We only consider overestimators, since the treatment for underestimators is similar. A linear overestimator is a linear function which provides an upper bound to a univariate function over an interval.

Definition 1 (Linear Overestimators). *Let g be a univariate function $\Re \to \Re$. A linear overestimator of g over the interval $[\underline{x}, \overline{x}]$ is a linear function $mx+b$ satisfying*

$$mx + b \geq g(x), \forall x \in [\underline{x}, \overline{x}].$$

In general, given a univariate function g, linear overestimators are obtained through tangent or secant lines. These are implicitly specified by two functions $f_m(\underline{x}, \overline{x}, g)$ and $f_b(\underline{x}, \overline{x}, g)$ respectively computing the slope m and the intercept b of the estimator.[1] For instance, the secant line for the function x^n (n even) is the linear overestimator $mx + b$ over $[\underline{x}, \overline{x}]$ specified by

$$m = \frac{\overline{x}^n - \underline{x}^n}{\overline{x} - \underline{x}} \qquad b = \frac{\overline{x}\underline{x}^n - \underline{x}\overline{x}^n}{\overline{x} - \underline{x}}$$

Unfortunately, given an underlying floating-point system $\mathcal{F}$ and a representation of f_m and f_b, the computation of these functions is subject to rounding errors and will produce the approximations $\tilde{m}$ and $\tilde{b}$. However, the linear function $\tilde{m}x + \tilde{b}$ is not guaranteed to be a linear estimator of g in $[\underline{x}, \overline{x}]$.

The main issue addressed in this paper is how to compute safe linear overestimators, i.e., linear overestimators $m^*x + b^*$ where m^* and b^* are floating-point numbers in the underlying floating-point system $\mathcal{F}$.

Definition 2 (Safe Linear Overestimators). *Let g be a univariate function $\Re \to \Re$. A* safe linear overestimator *of g over interval $[\underline{x}, \overline{x}]$ is a linear overestimator $m^*x + b^*$ for g over $[\underline{x}, \overline{x}]$ where $m^*, b^* \in \mathcal{F}$.*

[1] More precisely, they are specified by a representation (e.g., the text) of these two functions.

Notations In the following, we often abuse notation and use m and b to represent the functions f_m and f_b. The critical point to remember is that m and b cannot be computed exactly and may involve significant rounding errors. Also, given an expression e, we use $\lfloor e \rfloor$ and $\lceil e \rceil$ to denote the most precise lower and upper approximation of e at our disposal given $\mathcal{F}$ and the representation of e.[2]

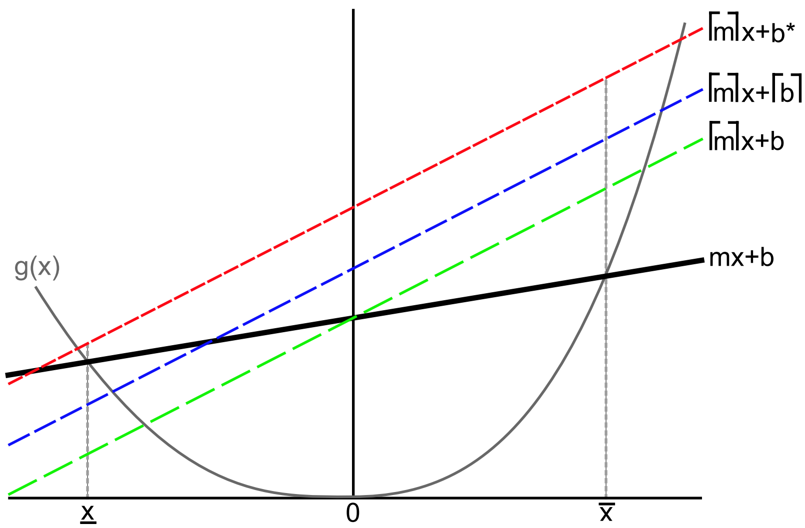

Fig. 1. Why $\lceil m \rceil x + \lceil b \rceil$ is not a Safe Linear Overestimator.

The Problem At first sight, it may seem that the problem of finding a safe linear overestimator $m^*x + b^*$ is trivial: simply choose the function $\lceil m \rceil x + \lceil b \rceil$, i.e., choose $m^* = \lceil m \rceil$ and $b^* = \lceil b \rceil$. Unfortunately, as shown in Figure 1, this is not correct. The figure shows g and its linear overestimator $mx + b$ over $[\underline{x}, \overline{x}]$. The estimator is correct in the $\Re^+$ region, but not in the $\Re^-$ region where the slope $\lceil m \rceil$ is too strong. Similarly, $\lfloor m \rfloor x + \lceil b \rceil$ is not a safe overestimator because its slope is too weak in the $\Re^+$ region. The value b^* must be chosen carefully when $m^* = \lceil m \rceil$ or $\lfloor m \rfloor$. The figure shows such a choice of b^*.

Tightness In addition to safety, one is generally interested in linear overestimators which are as tight as possible given $\mathcal{F}$ and the representation of f_m and f_b.

Definition 3 (Error of Linear Overestimators). *Let g be a univariate function $\Re \rightarrow \Re$. The error of a linear overestimator $mx + b$ for g over $[\underline{x}, \overline{x}]$ is given*

[2] Note the safety results presented in this paper hold even if the approximations are not the most precise.

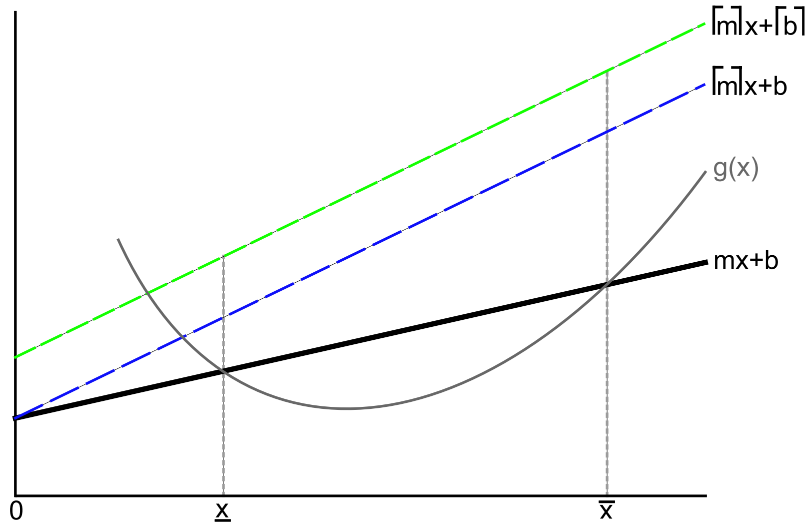

Fig. 2. A Safe Overestimator when $\underline{x} \geq 0$.

by

$$\int_{\underline{x}}^{\overline{x}} mx + b \;-\; g(x)\, dx.$$

Definition 4 (Tightness of Linear Overestimators). *Let g be a univariate function $\Re \rightarrow \Re$. Let l_1 and l_2 be two linear overestimators of g over $[\underline{x}, \overline{x}]$. l_1 is tighter than l_2 wrt g and $[\underline{x}, \overline{x}]$ if l_1 has a smaller error than l_2,*

3 Safe Linear Overestimators for Univariate Functions

This section describes two classes of safe overestimators. These estimators are derived from the linear overestimator $mx + b$. In other words, the goal is to find m^* and b^* in $\mathcal{F}$ such that

$$m^* x + b^* \geq mx + b,\ \forall x \in [\underline{x}, \overline{x}]. \tag{1}$$

As mentioned, the first class generalizes the results of Michel et al. [9] who gave safe estimators for x^2. The second class is entirely new and enjoys some nice theoretical and numerical properties.

3.1 A First Class of Safe Overestimators

To obtain a safe linear overestimator $m^* x + b^*$ from $mx + b$, there are only two reasonable choices for m^*: $\lfloor m \rfloor$ and $\lceil m \rceil$. Other choices would necessarily be less tight. We derive the safe overestimators for $\lceil m \rceil$ only, the derivation for $\lfloor m \rfloor$ being similar. The problem thus reduces to finding b^* such that

$$\lceil m \rceil x + b^* \geq mx + b. \tag{2}$$

Since $\lceil m \rceil \geq m$, it is sufficient to satisfy (2) at $x = \underline{x}$, which implies $b^* \geq b - (\lceil m \rceil - m)\underline{x}$. The overestimator can now be derived by a case analysis on the sign of $\underline{x}$. If $\underline{x} \leq 0$, we have

$$b - (\lceil m \rceil - m)\underline{x} \leq b - (\lceil m \rceil - \lfloor m \rfloor)\underline{x} = b - \mathtt{err}(m)\underline{x}$$

where $\mathtt{err}(m) = \lceil m \rceil - \lfloor m \rfloor$. Therefore, choosing $b^* = \lceil b - \mathtt{err}(m)\underline{x} \rceil$ satisfies (2). If $\underline{x} \geq 0$, it is sufficient to choose $b^* = \lceil b \rceil$, as shown in Figure 2. The following theorem summarizes the results.

Theorem 1 (Safe Linear Overestimators for Univariate Functions, Class 1). *Let g be a univariate function g and let $mx + b$ be a linear overestimator for g in $[\underline{x}, \overline{x}]$. We have that*

$$mx + b \leq \begin{cases} \lfloor m \rfloor x + \lceil b + \mathtt{err}(m)\overline{x} \rceil & \textit{if } \overline{x} \geq 0 \\ \lfloor m \rfloor x + \lceil b \rceil & \textit{if } \overline{x} \leq 0 \\ \lceil m \rceil x + \lceil b - \mathtt{err}(m)\underline{x} \rceil & \textit{if } \underline{x} \leq 0 \\ \lceil m \rceil x + \lceil b \rceil & \textit{if } \underline{x} \geq 0. \end{cases}$$

As a consequence, the four right-hand sides are safe linear overestimators for g in $[\underline{x}, \overline{x}]$ under the specified conditions.

Proof. We give the proofs for the first two cases. The proofs are similar for the symmetric cases. In the first case, we have

$$\begin{aligned}
&\lfloor m \rfloor x + \lceil b + \mathtt{err}(m)\overline{x} \rceil \\
&= \lfloor m \rfloor(\overline{x} - s) + \lceil b + \mathtt{err}(m)\overline{x} \rceil && \text{letting } x = \overline{x} - s \text{ with } s \geq 0 \\
&\geq \lfloor m \rfloor(\overline{x} - s) + b + \mathtt{err}(m)\overline{x} \\
&= \lfloor m \rfloor(\overline{x} - s) + b + (\lceil m \rceil - \lfloor m \rfloor)\overline{x} \\
&= \lceil m \rceil \overline{x} - \lfloor m \rfloor s + b \\
&\geq m\overline{x} - \lfloor m \rfloor s + b && \lceil m \rceil \overline{x} \geq m\overline{x} \text{ since } \overline{x} \geq 0 \\
&\geq m\overline{x} - ms + b && \lfloor m \rfloor s \leq ms \text{ since } s \geq 0 \\
&= mx + b
\end{aligned}$$

In the second case, since $x \leq \overline{x} \leq 0$, we have that $mx \leq \lfloor m \rfloor x$ and hence $mx + b \leq \lfloor m \rfloor x + \lceil b \rceil$. $\square$

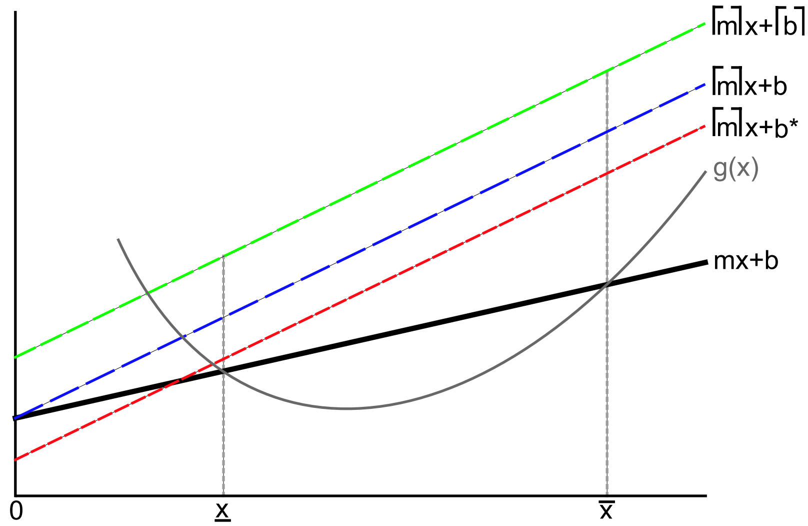

Fig. 3. Choosing $b^* = \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ which is almost always tighter than $\lceil m \rceil x + \lceil b \rceil$.

3.2 A Second Class of Safe Overestimators

In general, the overestimators used in global optimization are either secant lines of g (as in Figure 1) or tangent lines to g at $\underline{x}$ or $\overline{x}$ (see, for instance, [5]). As a consequence, we have that $g(\underline{x}) = m\underline{x} + b$ and/or $g(\overline{x}) = m\overline{x} + b$. In these circumstances, it is possible to find a b^* satisfying (2) which does not depend on the sign of $\underline{x}$. Assume that $g(\underline{x}) = m\underline{x} + b$. Since b^* must satisfy $\lceil m \rceil \underline{x} + b^* \geq m\underline{x} + b = g(\underline{x})$, it follows that $b^* \geq g(\underline{x}) - \lceil m \rceil \underline{x}$. Choosing $b^* = \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ also satisfies (2). This choice for b^* enjoys some nice theoretical and experimental properties as detailed in Section 4. Figure 3 illustrates this situation.

Theorem 2 (Safe Linear Overestimators of Univariate Functions, Class 2). *Let g be a univariate function g and let $mx + b$ be a linear overestimator for g in $[\underline{x}, \overline{x}]$. We have that*

$$mx + b \leq \begin{cases} \lfloor m \rfloor x + \lceil g(\overline{x}) - \lfloor m \rfloor \overline{x} \rceil & \text{if } \ g(\overline{x}) = m\overline{x} + b \\ \lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil & \text{if } \ g(\underline{x}) = m\underline{x} + b \end{cases}$$

As a consequence, the right-hand sides are safe linear overestimators for g in $[\underline{x}, \overline{x}]$ under the specified conditions.

Proof. We give the proof for the first case. The proof is similar for the symmetric case. We have

$$\begin{aligned}
&\lfloor m \rfloor x + \lceil g(\overline{x}) - \lfloor m \rfloor \overline{x} \rceil && \\
&= \lfloor m \rfloor (\overline{x} - s) + \lceil g(\overline{x}) - \lfloor m \rfloor \overline{x} \rceil && \text{letting } x = \overline{x} - s \text{ with } s \geq 0 \\
&\geq \lfloor m \rfloor (\overline{x} - s) + g(\overline{x}) - \lfloor m \rfloor \overline{x} && \\
&= m\overline{x} + b - \lfloor m \rfloor s && \text{since } g(\overline{x}) = m\overline{x} + b \\
&\geq m\overline{x} + b - ms && \lfloor m \rfloor s \leq ms \text{ since } s \geq 0 \\
&= mx + b. &&
\end{aligned}$$

□

4 Tightness of Safe Linear Overestimators

Theorems 1 and 2 provide us with six safe overestimators of a univariate function. Several of the conditions for these estimators are not mutually exclusive and it is natural to study their relative tightness. Of course, it is possible to use all applicable safe estimators in the linear relaxation, but this may be undesirable for numerical and efficiency reasons. This section presents theoretical and experimental results on the tightness of the estimators.

4.1 Theoretical Results on Class 1 Estimators

This section studies the tightness of class-1 estimators. We first compare the estimators $\lfloor m \rfloor x + \lceil b + \texttt{err}(m)\overline{x} \rceil$ and $\lceil m \rceil x + \lceil b - \texttt{err}(m)\underline{x} \rceil$ which are both applicable when $0 \in [\underline{x}, \overline{x}]$. The result shows which estimators to choose according to the magnitude of $\underline{x}$ and $\overline{x}$. Figure 4 illustrates this.

Theorem 3 (Tightness of Class 1 Safe Linear Overestimators when $0 \in [\underline{x}, \overline{x}]$). *Let g be a univariate function g, $mx+b$ be a linear overestimator for g in $[\underline{x}, \overline{x}]$, and $\underline{x} < 0$ and $\overline{x} > 0$. The safe linear overestimator $\lfloor m \rfloor x + \lceil b + \texttt{err}(m)\overline{x} \rceil$ is tighter than the safe linear overestimator $\lceil m \rceil x + \lceil b - \texttt{err}(m)\underline{x} \rceil$ if $|\overline{x}| < |\underline{x}|$. Similarly, the safe linear overestimator $\lceil m \rceil x + \lceil b - \texttt{err}(m)\underline{x} \rceil$ is tighter than the safe linear overestimator $\lfloor m \rfloor x + \lceil b + \texttt{err}(m)\overline{x} \rceil$ if $|\underline{x}| < |\overline{x}|$.*

Proof. To compare the estimators $\lfloor m \rfloor x + \lceil b + \texttt{err}(m)\overline{x} \rceil$ and $\lceil m \rceil x + \lceil b - \texttt{err}(m)\underline{x} \rceil$, we compare the relative tightness of the slightly tighter estimators $\lfloor m \rfloor x + b + \texttt{err}(m)\overline{x}$ and $\lceil m \rceil x + b - \texttt{err}(m)\underline{x}$. Their tightness is easier to determine and approximates well the actual relative tightness, since the rounding errors in computing $\lceil b + \texttt{err}(m)\overline{x} \rceil$ and $\lceil b - \texttt{err}(m)\underline{x} \rceil$ are comparable. First consider the error of $\lfloor m \rfloor x + b + \texttt{err}(m)\overline{x}$:

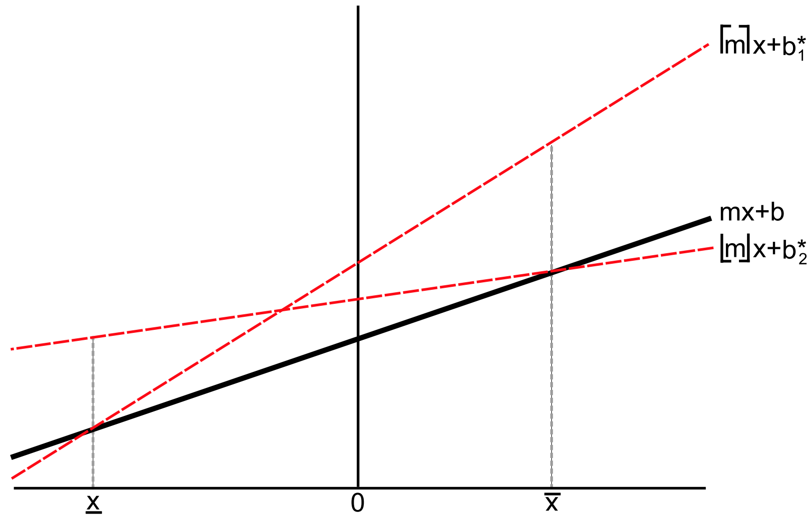

Fig. 4. Finding the optimal floating point representation. In this case $\lfloor m \rfloor x + b_2^*$ is a tighter estimator than $\lceil m \rceil x + b_1^*$.

$$
\begin{aligned}
E_1 &= \int_{\underline{x}}^{\overline{x}} \lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x} - g(x)\mathrm{d}x \\
&= \int_{\underline{x}}^{\overline{x}} \lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x} - (mx + b) + mx + b - g(x)\mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{\underline{x}}^{\overline{x}} (\lfloor m \rfloor - m)x + \mathtt{err}(m)\overline{x}\mathrm{d}x + \underbrace{\int_{\underline{x}}^{\overline{x}} mx + b - g(x)\mathrm{d}x}_{E} \\
&= (\overline{x} - \underline{x}) \left[\overline{x} \left(\lceil m \rceil - \frac{1}{2}\lfloor m \rfloor - \frac{1}{2}m \right) + \underline{x} \left(\frac{1}{2}\lfloor m \rfloor - \frac{1}{2}m \right) \right] + E
\end{aligned}
$$

The error of $\lceil m \rceil x + b - \mathtt{err}(m)\underline{x}$ is similarly given by:

$$
\begin{aligned}
E_2 &= \int_{\underline{x}}^{\overline{x}} \lceil m \rceil x + b - \mathtt{err}(m)\underline{x} - g(x)\mathrm{d}x \\
&= (\overline{x} - \underline{x}) \left[\overline{x} \left(\frac{1}{2}\lceil m \rceil - \frac{1}{2}m \right) + \underline{x} \left(\lfloor m \rfloor - \frac{1}{2}\lceil m \rceil - \frac{1}{2}m \right) \right] + E
\end{aligned}
$$

Estimator $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$ is tighter than estimator $\lceil m \rceil x + b - \mathtt{err}(m)\underline{x}$ when $E_1 < E_2$, i.e., when $|\overline{x}| < |\underline{x}|$. Similarly, estimator $\lceil m \rceil x + b - \mathtt{err}(m)\underline{x}$ is tighter than estimator $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$ when $|\underline{x}| < |\overline{x}|$. □

When $\underline{x} \geq 0$, it is interesting to compare estimators $\lfloor m \rfloor x + \lceil b + \mathtt{err}(m)\overline{x} \rceil$ with $\lceil m \rceil x + \lceil b \rceil$ and $\lfloor m \rfloor x + \lceil b \rceil$ with $\lceil m \rceil x + \lceil b - \mathtt{err}(m)\underline{x} \rceil$.

Theorem 4 (Tightness of Class 1 Safe Linear Overestimators when $0 \notin [\underline{x}, \overline{x}]$). *Let g be a univariate function g and $mx+b$ be a linear overestimator for g in $[\underline{x}, \overline{x}]$. When $\underline{x} \geq 0$, $\lceil m \rceil x + b$ is tighter than $\lfloor m \rfloor x + \lceil b + \mathtt{err}(m)\overline{x} \rceil$. When $\overline{x} \leq 0$, $\lfloor m \rfloor x + \lceil b \rceil$ is tighter than $\lceil m \rceil x + \lceil b - \mathtt{err}(m)\underline{x} \rceil$.*

Proof. Consider the slightly tighter and easily comparable estimators $\lceil m \rceil x + b$ and $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$. The error of $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$ is:

$$\begin{aligned} E_1 &= \int_{\underline{x}}^{\overline{x}} \lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x} - g(x)\mathrm{d}x \\ &= \frac{1}{2}(\overline{x} - \underline{x})\left[\overline{x}\left(2\lceil m \rceil - \lfloor m \rfloor - m\right) + \underline{x}\left(\lfloor m \rfloor - m\right)\right] + E \end{aligned}$$

where E is the error in $mx + b$. Likewise the error of $\lceil m \rceil x + b$ is:

$$\begin{aligned} E_2 &= \int_{\underline{x}}^{\overline{x}} \lceil m \rceil x + b - g(x)\mathrm{d}x \\ &= \frac{1}{2}(\overline{x} - \underline{x})(\lceil m \rceil - m)(\underline{x} + \overline{x}) + E \end{aligned}$$

$\lceil m \rceil x + b$ is tighter than $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$ when $E_2 < E_1$ which reduces to $\underline{x} < \overline{x}$. Since this condition is always met, $\lceil m \rceil x + b$ is always tighter than $\lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x}$. Similarly, $\lfloor m \rfloor x + \lceil b \rceil$ is tighter than $\lceil m \rceil x + \lceil b - \mathtt{err}(m)\underline{x} \rceil$ when $\overline{x} \leq 0$. □

Theorems 3 and 4 generalize and provide the theoretical justification for the heuristic used by Michel et al. [9].

4.2 Theoretical Results on Class 2 Estimators

We now study the tightness of Class-2 estimators. The next theorem compares the "real" counterparts of the two class-2 operators.

Theorem 5 (Tightness of Class 2 Safe Linear Overestimators). *Let g be a univariate function with linear overestimator $mx + b$ over $[\underline{x}, \overline{x}]$ such that $g(\underline{x}) = m\underline{x} + b$, $g(\overline{x}) = m\overline{x} + b$. $\lfloor m \rfloor x + g(\overline{x}) - \lfloor m \rfloor \overline{x}$ is a tighter estimator than $\lceil m \rceil x + g(\underline{x}) - \lceil m \rceil \underline{x}$ when $m - \lfloor m \rfloor < \lceil m \rceil - m$.*

Proof. The error of the $g(\overline{x}) = m\overline{x} + b$. $\lfloor m \rfloor x + g(\overline{x}) - \lfloor m \rfloor \overline{x}$ is given by

$$\begin{aligned} E_1 &= \int_{\underline{x}}^{\overline{x}} \lfloor m \rfloor x + g(\overline{x}) - \lfloor m \rfloor \overline{x} - g(x) \\ &= \frac{1}{2}(\overline{x} - \underline{x})^2(m - \lfloor m \rfloor) + E \end{aligned}$$

where E is the error of $mx + b$. Likewise the error in $\lceil m \rceil x + g(\underline{x}) - \lceil m \rceil \underline{x}$ is

$$E_2 = \frac{1}{2}(\overline{x} - \underline{x})^2(\lceil m \rceil - m) + E.$$

The $\lfloor m \rfloor$ formulation is tighter when $E_1 < E_2$ which reduces to $m - \lfloor m \rfloor < \lceil m \rceil - m$, i.e., when $\lfloor m \rfloor$ is a better approximation of m than $\lceil m \rceil$. □

Of course, this result is not useful in practice, since m is not known and there is no way to evaluate the condition stated in Theorem 5. The theorem only considers the "real" counterparts of the estimators, i.e., it ignores the rounding errors in the actual evaluation of the operators. In other words, since these operators can be rewritten as $\lceil \lfloor m \rfloor (x - \overline{x}) + g(\overline{x}) \rceil$ and $\lceil \lceil m \rceil (x - \underline{x}) + g(\underline{x}) \rceil$, Theorem 5 applies to the Class-2 estimators whenever the rounding errors in these two terms are similar which in turn requires that $g(\underline{x}) \sim g(\overline{x})$. Fortunately, the next section, which relates both classes, gives a criterion to choose between them.

4.3 Theoretical Results on Class 1 and Class 2 Estimators

This section compares the Class 1 and Class 2 overestimators. Its main result shows that class-2 estimators are always theoretically tighter than the corresponding optimal class-1 estimators. The theorems are given for the $\lceil m \rceil$ estimators, but similar results hold for the $\lfloor m \rfloor$ estimators.

Theorem 6 (Relative Tightness of Class 1 and Class 2 Safe Linear Overestimators). *Let g be a univariate function with linear overestimator $mx + b$ over $[\underline{x}, \overline{x}]$ such that $g(\underline{x}) = m\underline{x} + b$. The class-2 estimator ($\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$) is always tighter than the optimal (using the rules given in Theorems 3 and 4) class-1 estimator when $|\overline{x}| \geq |\underline{x}|$ (modulo rounding errors).*

Proof. We prove the result modulo rounding errors. When $\underline{x} \leq 0$, we have

$$\begin{aligned}
&\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil \\
&\leq \lceil m \rceil x + \lceil g(\underline{x}) - (m + \mathtt{err}(m))\underline{x} \rceil \\
&\leq \lceil m \rceil x + \lceil b - \mathtt{err}(m)\underline{x} \rceil
\end{aligned}$$

which is the class-1 estimator when $\underline{x} \leq 0$. When $\underline{x} \geq 0$, we have

$$\begin{aligned}
&\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil \\
&\leq \lceil m \rceil x + \lceil g(\underline{x}) - m\underline{x} \rceil \\
&\leq \lceil m \rceil x + \lceil b \rceil
\end{aligned}$$

which is the class-1 estimator when $\underline{x} \geq 0$. □

Theorem 6 is interesting for many reasons. First, although it abstracts the rounding errors, it should hold in practice since the rounding errors in $g(\underline{x})$ should be smaller than, or similar to, those in b. (There are of course other terms but the errors cannot be radically different in the remaining parts). The experimental results in Section 4.4 confirms this. Second, it provides intuition as to why class-2 estimators are tighter than class-1 estimators. A class-1 operator systematically accounts for an error $\texttt{err}(m) = \lceil m \rceil - \lfloor m \rfloor$ in the slope, while a class-2 estimator only adds an error $\lceil m \rceil - m$ (resp. $m - \lfloor m \rfloor$). Third, since class-2 operators can be viewed as tighter versions of the class-1 operators, similar criteria can be applied to choose between them, solving the problem left open in the previous section.

For completeness, Appendix A compares the class-2 safe linear overestimator $\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ with the class-1 estimators that uses $m^* = \lfloor m \rfloor$. This result is not useful in practice when the class-2 operator using $\lfloor m \rfloor$ is available (which is always the case for secant lines for instance). In this case, one should choose the other class-2 estimator which is guaranteed to be tighter theoretically. However, the result sheds some light on the relationships between the two classes of estimators and indicate that, in general, class-2 estimators should be preferred.

4.4 Numerical Results

We now compare numerically class-1 and class-2 estimators to confirm the findings of Theorem 6. More precisely, we compare $\lfloor m \rfloor x + \lceil g(\overline{x}) - \lfloor m \rfloor \overline{x} \rceil$ with $\lfloor m \rfloor x + \lceil b + \texttt{err}(m)\overline{x} \rceil$ and $\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ with $\lceil m \rceil x + \lceil b - \texttt{err}(m)\underline{x} \rceil$ numerically for $0 \in [\underline{x}, \overline{x}]$ using the above heuristics to choose between the pairs. We use even powers for comparison purposes for which linear overestimators are given as follows.

$$g(x) = x^n \leq \underbrace{\frac{\overline{x}^n - \underline{x}^n}{\overline{x} - \underline{x}}}_{m} x + \underbrace{\frac{\overline{x}\underline{x}^n - \underline{x}\overline{x}^n}{\overline{x} - \underline{x}}}_{b}, \; n \text{ even}, \; x \in [\underline{x}, \overline{x}] \tag{3}$$

This general term can be simplified using $m = \underline{x} + \overline{x}$ and $b = -\underline{x}\overline{x}$ when $n = 2$. Moreover, since (3) is a secant of x^n, $g(x) = mx + b$ at both $x = \underline{x}$ and $x = \overline{x}$.

Given the theoretical results, it is easy to derive a set of numerical experiments. When $|\underline{x}| \geq |\overline{x}|$, we only compare the intercepts $\lceil g(\overline{x}) - \lfloor m \rfloor \overline{x} \rceil$ and $\lceil b + \texttt{err}(m)\overline{x} \rceil$, using the estimators with the same slopes. The smallest intercept gives the tightest estimator. Likewise when $|\underline{x}| < |\overline{x}|$, we compare the intercepts $\lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ and $\lceil b - \texttt{err}(m)\underline{x} \rceil$.

Figures 5 and 6 depict the experimental results. To compute b, our numerical results use the specialized form for $n = 2$ (Figure 5) and the general form (Figure 6) otherwise. Random values were generated for $\underline{x}$ and $\overline{x}$ in a wide range of values. The results were collected region by region. We computed the percentage of cases where class-2 estimators were strictly tighter than class-1 estimators (%C21) and vice-versa (%C12). The figures report the difference (%C21 - %C12). For the quadratic case, Figure 5 shows that class-2 estimators are very often tighter than class-1 operators. Typically, class-2 estimators are tighter in about 40-50% of the cases, while class-1 estimators are tighter in about 0-10% of the cases (they have equal tightness in the remaining cases). It is only when the two bounds are about the same size that class-1 improves over class-2 in about 40-50% of the

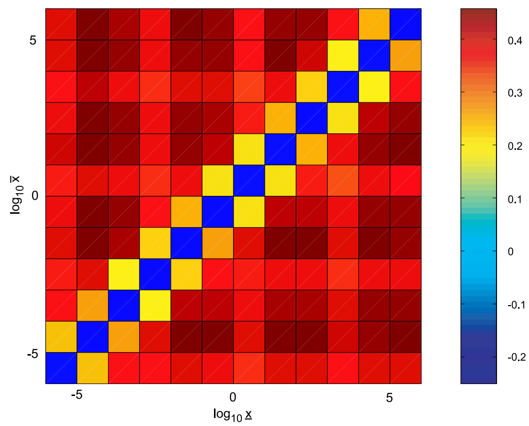

Fig. 5. Numerical Results for x^2.

cases. Figure 6 depicts the results for n^4 to n^{10}. The improvements of class-2 estimators here is striking. Class-2 estimators improve over class-1 estimators in more than 99% of the cases. Of course, this is not surprising in light of the proof of Theorem 6 since the errors in the slope multiplying $\underline{x}$ or $\overline{x}$ become larger for class-1 estimators and the functions f_m and f_b are also more complex than x^n when $n > 2$, leading to more pronounced rounding errors. Note also the improvement in tightness is proportional to the magnitude of the bounds. In summary, the experimental results confirm the theoretical results and clearly demonstrate the value of class-2 estimators.

5 Safe Linear Overestimators for Multivariate Functions

We now turn our attention to safe linear overestimators for multivariate functions. Multivariate functions frequently appear in global optimization. They can be estimated using a hyperplane in n dimensions. For example, a linear overestimator for the bilinear term xy (e.g., [5, 11]) is given by two planes:

$$xy \leq \min\{\underline{y}x + \overline{x}y - \overline{x}\underline{y}, \overline{y}x + \underline{x}y - \underline{x}\overline{y}\},\ (x, y) \in [\underline{x}, \overline{x}] \times [\underline{y}, \overline{y}] \quad (4)$$

Since slopes of the two planes given in (4) are floating-point numbers, safe estimators for this case are given simply by rounding up the intercepts $-\overline{x}\underline{y}$ and $-\underline{x}\overline{y}$. The overestimator for the term $\frac{x}{y}$ used by [11, 16] have slopes that are nonlinear functions of floating point numbers:

$$\frac{x}{y} \leq \min\{\frac{1}{\underline{y}\overline{y}}(\overline{y}x - \underline{x}y + \underline{x}\underline{y}), \frac{1}{\underline{y}\overline{y}}(\underline{y}x - \overline{x}y + \overline{x}\overline{y})\}(x, y) \in [\underline{x}, \overline{x}] \times [\underline{y}, \overline{y}], \underline{y} > 0 \quad (5)$$

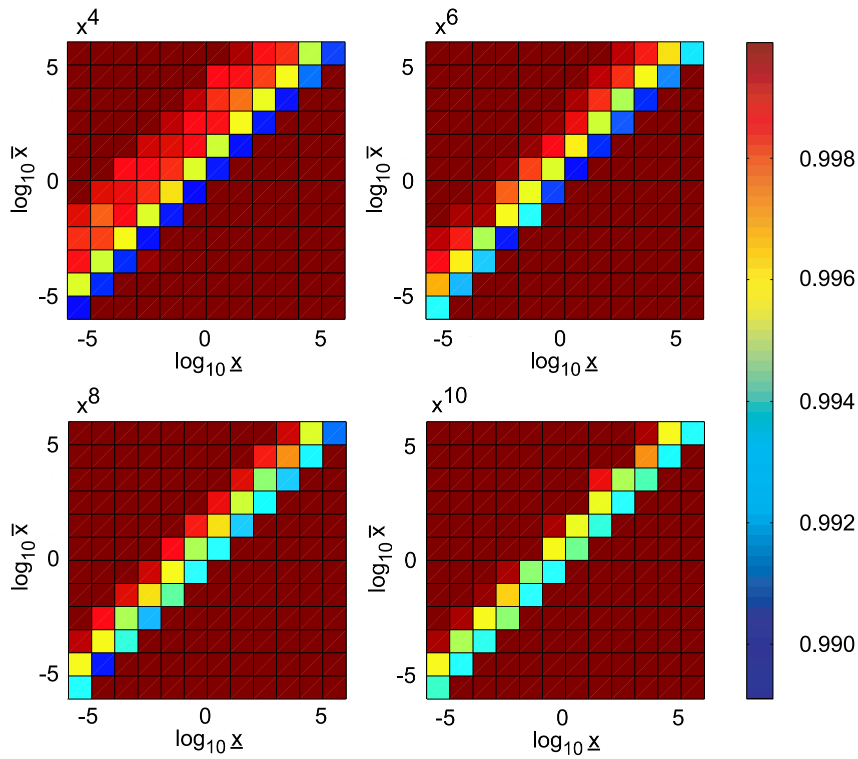

Fig. 6. Numerical Results for x^4 to x^{10}.

In order to represent (5) with floating point numbers, special care must be taken in order to satisfy the two dimensional version of (1). More generally, estimators for multivariate functions can be obtained through estimators for factorable functions (see, for instance, [8]).

The above discussion indicates that it is critical in practice to generalize our results to multivariate functions. The problem can be formalized as follows (for overestimators): given an n-dimensional hyperplane overestimating an n-variate function $g(\mathbf{x}) : \Re^n \rightarrow \Re$, $\mathbf{x} = (x_1, \ldots, x_n)$, over $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$, the goal is to find $m_1^*, \ldots, m_n^*, b^* \in \mathcal{F}$ such that:

$$\sum_{i=1}^{n} m_i^* x_i + b^* \geq \sum_{i=1}^{n} m_i x_i + b \geq g(\mathbf{x}),\ \forall \mathbf{x} \in [\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n] \tag{6}$$

The first key result in this section is to show that safe linear estimators for multivariate functions can be derived naturally by combining univariate linear estimators. In other words, the result makes it possible to consider each variable independently in the estimator

$$\sum_{i=1}^{n} m_i x_i + b,$$

replace m_i and b by their safe counterparts m_i^* and b_i^*, and combine all the individual coefficients into a safe estimator for the multivariate function. One of the interesting aspects of this result is the ability to factor b out from the b_i^*s to obtain a tight estimator.

Theorem 7 (Safe Linear Overestimators for Multivariate Functions). *Let g be an n-variate function and let $\sum_{i=1}^{n} m_i x_i + b$ be an overestimator for g in $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$. Let $m_i^* x_i + b_i^*$ be a safe linear overestimator of $m_i x_i + b$ in $[\underline{x}_i, \overline{x}_i]$ and $\delta_i^* = \lceil b_i^* - b \rceil$, $1 \leq i \leq n$. Then, the hyperplane*

$$\sum_{i=1}^{n} m_i^* x_i + \lceil b + \sum_{i=1}^{n} \delta_i^* \rceil$$

is a safe linear overestimator for g in $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$.

Proof. We show that

$$\sum_{i=1}^{n} m_i^* x_i + \lceil b + \sum_{i=1}^{n} \delta_i^* \rceil \geq \sum_{i=1}^{n} m_i x_i + b.$$

We have

$$\begin{aligned}
\sum_{i=1}^{n} m_i^* x_i + \lceil b + \sum_{i=1}^{n} \delta_i^* \rceil &\geq \sum_{i=1}^{n} m_i^* x_i + b + \sum_{i=1}^{n} \delta_i^* \\
&\geq \sum_{i=1}^{n} (m_i^* x_i + b_i^* - b) + b \\
&= \sum_{i=1}^{n} (m_i^* x_i + b_i^*) - (n-1)b \\
&\geq \sum_{i=1}^{n} (m_i x_i + b) - (n-1)b \\
&= \sum_{i=1}^{n} m_i x_i + b. \quad \square
\end{aligned}$$

Note that class-1 estimators trivially satisfy the requirements of this theorem. As a consequence, the theorem gives an elegant and simple way to obtain safe overestimators for multivariate functions. (A similar result can be derived for underestimators).

5.1 Class-1 Safe Multivariate Linear Estimators

We now investigate the tightness of class-1 safe multivariate estimators derived using Theorem 7. The class-1 safe multivariate estimators are derived by combining class-1 univariate estimators. Our first result shows that class-1 multivariate estimators derived using Theorem 7 are as tight as possible for a given choice of rounding direction for each dimension (i.e., $\lfloor m_i \rfloor$ or $\lceil m_i \rceil$).

Theorem 8 (Tightness of Class-1 Safe Multivariate Overestimators). *Class-1 multivariate overestimators derived using Theorem 7 are as tight as possible for a given choice of rounding direction for each variable (i.e., $\lfloor m_i \rfloor$ or $\lceil m_i \rceil$).*

Proof. The safe overestimating hyperplane given by Theorem 7 can be written as:

$$\sum_{i \in I \cup K} \lceil m_i \rceil x_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor x_j + \lceil b - \sum_{k \in K} \mathtt{err}(m_k)\underline{x}_k + \sum_{l \in L} \mathtt{err}(m_l)\overline{x}_l \rceil$$

where sets I, J, K and L partition $\{1, \ldots, n\}$. Using the same choices for $\lceil m \rceil$ and $\lfloor m \rfloor$ defined by this partition, we now calculate b^* such that $\sum_{i \in I \cup K} \lceil m_i \rceil x_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor x_j + b^* \geq \sum_{i=1}^{n} m_i x_i + b$. It is sufficient to satisfy this inequality at $(\tilde{x}_1, \ldots, \tilde{x}_n)$ where

$$\tilde{x}_i = \begin{cases} \overline{x}_i & : i \in J \cup L \\ \underline{x}_i & : \text{otherwise} \end{cases}$$

due to the combination of $\lfloor m \rfloor$ and $\lceil m \rceil$ given by the partition. b^* must satisfy

$$\sum_{i \in I \cup K} \lceil m_i \rceil \underline{x}_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor \overline{x}_j + b^* \geq \sum_{i \in I \cup K} m_i \underline{x}_i + \sum_{j \in J \cup L} m_j \overline{x}_j + b.$$

By a case analysis on the sign of $\tilde{x}_i$ for each i:

$$\begin{aligned} & \sum_{i \in I \cup K} m_i \underline{x}_i + \sum_{j \in J \cup L} m_j \overline{x}_j + b \\ & \leq \sum_{i \in I} \lceil m_i \rceil \underline{x}_i + \sum_{j \in J} \lfloor m_j \rfloor \overline{x}_j + \sum_{k \in K} \lfloor m_k \rfloor \underline{x}_k + \sum_{l \in L} \lceil m_l \rceil \overline{x}_l + b \end{aligned}$$

Therefore,

$$\begin{aligned} & \sum_{i \in I \cup K} \lceil m_i \rceil \underline{x}_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor \overline{x}_j + b^* \\ & \geq \sum_{i \in I} \lceil m_i \rceil \underline{x}_i + \sum_{j \in J} \lfloor m_j \rfloor \overline{x}_j + \sum_{k \in K} \lfloor m_k \rfloor \underline{x}_k + \sum_{l \in L} \lceil m_l \rceil \overline{x}_l + b. \end{aligned}$$

Collecting terms: $b^* \geq b - \sum_{k \in K} \mathtt{err}(m_k)\underline{x}_k + \sum_{l \in L} \mathtt{err}(m_l)\overline{x}_l$. Correctly rounding this results in the same hyperplane as constructed by the theorem. □

It remains to choose appropriate rounding directions for each variable. The next theorem shows that the combination of tight class-1 univariate estimators gives a tight class-1 multivariate estimator.

Theorem 9 (Optimality of Class-1 Safe Multivariate Overestimators). *A multivariate class-1 safe estimators derived using Theorem 7 by choosing optimal class-1 univariate safe estimators for each variable is an optimal class-1 multivariate safe estimator.*

Proof. Consider the hyperplane:

$$H = \sum_{i \in I \cup K} \lceil m_i \rceil x_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor x_j + \lceil b - \sum_{k \in K} \mathtt{err}(m_k) \underline{x}_k + \sum_{l \in L} \mathtt{err}(m_l) \overline{x}_l \rceil \quad (7)$$

where $I = \{i : \underline{x}_i \geq 0\}$, $J = \{j : \overline{x}_j \leq 0\}$, $K = \{k : 0 \in (\underline{x}_k, \overline{x}_k), |\underline{x}_k| < |\overline{x}_k|\}$ and $L = \{l : 0 \in (\underline{x}_l, \overline{x}_l), |\underline{x}_l| \geq |\overline{x}_l|\}$. By theorem 7, H is safe. By theorem 8, H is as tight as possible given the choices of $\lceil m \rceil$ and $\lfloor m \rfloor$ defined by the partition. Consider the error between a slightly tighter version of (7), given by an unrounded intercept, and $\sum_{i=1}^{n} m_i x_i + b$. The remaining error $(\int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \sum_{i=1}^{n} m_i x_i + b - g)$ is constant over all 2^n possible safe class-1 overestimating hyperplanes. The error is defined by the natural extension of definition 3 to higher dimensions:

$$\begin{aligned} E = &\int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \left\{ \sum_{i \in I \cup K} \lceil m_i \rceil x_i + \sum_{j \in J \cup L} \lfloor m_j \rfloor x_j + b - \sum_{k \in K} \mathtt{err}(m_k) \underline{x}_k \right. \\ &\qquad\qquad \left. + \sum_{l \in L} \mathtt{err}(m_l) \overline{x}_l - \left(\sum_{i=1}^{n} m_i x_i + b \right) \right\} \\ = &\int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \left\{ \sum_{i \in I \cup K} (\lceil m_i \rceil - m_i) x_i + \sum_{j \in J \cup L} (\lfloor m_j \rfloor - m_j) x_j \right. \\ &\qquad\qquad \left. - \sum_{k \in K} \mathtt{err}(m_k) \underline{x}_k + \sum_{l \in L} \mathtt{err}(m_l) \overline{x}_l \right\} \end{aligned}$$

Using the integrals

$$\begin{aligned} \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \; 1 &= \prod_{j=1}^{n} (\overline{x}_j - \underline{x}_j) = P \\ \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \; x_i &= \frac{1}{2} (\overline{x}_i^2 - \underline{x}_i^2) \prod_{j \neq i} (\overline{x}_j - \underline{x}_j) \\ &= \frac{1}{2} (\overline{x}_i + \underline{x}_i) P \end{aligned}$$

the error is rewritten as:

$$
\begin{aligned}
E = P & \left\{ \sum_{i \in I \cup K} \frac{1}{2}(\lceil m_i \rceil - m_i)(\overline{x}_i + \underline{x}_i) + \sum_{j \in J \cup L} \frac{1}{2}(\lfloor m_j \rfloor - m_j)(\overline{x}_i + \underline{x}_i) \right. \\
& \left. - \sum_{k \in K} \mathtt{err}(m_k)\underline{x}_k + \sum_{l \in L} \mathtt{err}(m_l)\overline{x}_l \right\} \\
= \frac{P}{2} & \left\{ \sum_{i \in I}(\lceil m_i \rceil - m_i)(\overline{x}_i + \underline{x}_i) + \sum_{j \in J}(\lfloor m_j \rfloor - m_j)(\overline{x}_i + \underline{x}_i) \right. \\
& + \sum_{k \in K}\{(\lceil m_i \rceil - m_i)\overline{x}_i + (2\lfloor m_i \rfloor - \lceil m_i \rceil - m_i)\underline{x}_i\} \qquad (8) \\
& \left. + \sum_{l \in L}\{(\lfloor m_i \rfloor - m_i)\underline{x}_i + (2\lceil m_i \rceil - \lfloor m_i \rfloor - m_i)\overline{x}_i\} \right\}
\end{aligned}
$$

Define the type of a variable as the set (one of I, J, K, or L) to which it belongs. We show that another partition, $\{I',\ J',\ K',\ L'\}$, does not define a tighter hyperplane, H'. Since the error given by (8) is linear in the type of variable, we can consider each variable independently. H and H' can differ in any of the following four ways while remaining safe:

- A variable, x_i, in set I can be moved to set L producing $I' = I \backslash \{x_i\}$ and $L' = L \cup \{x_i\}$. This will add the term $P \cdot \mathtt{err}(m_l)\overline{x}_l \geq 0$ to the error, thereby increasing the total error. H remains tighter than H'.
- A variable, x_i, in set J can be moved to set K producing $J' = J \backslash \{x_i\}$ and $K' = K \cup \{x_i\}$. This will add the term $-P \cdot \mathtt{err}(m_l)\underline{x}_l \geq 0$ to the error, thereby increasing the total error. H remains tighter than H'.
- A variable, x_i in set K can move to set L producing $K' = K \backslash \{x_i\}$ and $L' = L \cup \{x_i\}$. The difference in error is:

$$
\begin{aligned}
& \frac{P}{2}\{(\lfloor m_i \rfloor - m_i)\underline{x}_i + (2\lceil m_i \rceil - \lfloor m_i \rfloor - m_i)\overline{x}_i \\
& \quad - (\lceil m_i \rceil - m_i)\overline{x}_i - (2\lfloor m_i \rfloor - \lceil m_i \rceil - m_i)\underline{x}_i\} \\
& \quad = P \cdot \mathtt{err}(m)\{\overline{x}_i + \underline{x}_i\} \\
& \quad \geq 0 \ \text{ since the variable in } K \text{ satisfies } |\overline{x}| > |\underline{x}|.
\end{aligned}
$$

  Since the error increases with this change, H remains tighter than H'.
- Likewise, a variable moving from set L to set K increases the error of the overestimating hyperplane.

By the above arguments, the hyperplane defined by theorem 7 created by combining tight safe class-1 univariate overestimators is the tightest class-1 overestimator for an n-variate function. □

5.2 Class-2 Safe Multivariate Linear Estimators

The definition of class-2 estimators for multivariate functions is a direct extension of those for univariate functions. Notice that for the linear fractional term x/y, the overestimator $(\overline{y}x - \underline{x}y + \underline{xy})/\underline{y}\overline{y}$ is a plane through the points $(\underline{x}, \underline{y})$, $(\underline{x}, \overline{y})$, $(\overline{x}, \underline{y})$. As a result, a class-2 estimator can be defined at any of these points. In general, a purely convex n-variate function will have a linear overestimator passing through at least $n+1$ points (the secant hyperplane). A purely concave function will have a linear overestimator passing through one point (the tangent hyperplane). This information is used to define a safe overestimating hyperplane:

Theorem 10 (Class 2 Safe Multivariate Overestimators). *Let g be an n-variate function with linear overestimator $\sum_{i=1}^n m_i x_i + b$ over $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$ such that $g(\tilde{x}_1, \ldots, \tilde{x}_n) = \sum_{i=1}^n m_i \tilde{x}_i + b$ where $\tilde{x}_i \in \{\underline{x}_i, \overline{x}_i\}$. Let $M = \{i | \tilde{x}_i = \underline{x}_i\}$ and $N = \{i | \tilde{x}_i = \overline{x}_i\}$. $\sum_{i \in M} \lceil m_i \rceil x_i + \sum_{i \in N} \lfloor m_i \rfloor x_i + \lceil g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i \in M} \lceil m_i \rceil \underline{x}_i - \sum_{i \in N} \lfloor m_i \rfloor \overline{x}_i \rceil$ is a safe linear overestimator.*

Proof. The proof follows the same format as for the proof of Theorem 7:

$$
\begin{aligned}
&\sum_{i \in M} \lceil m_i \rceil x_i + \sum_{i \in N} \lfloor m_i \rfloor x_i + \lceil g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i \in M} \lceil m_i \rceil \underline{x}_i - \sum_{i \in N} \lfloor m_i \rfloor \overline{x}_i \rceil \\
\geq\ & \sum_{i \in M} \lceil m_i \rceil (\underline{x}_i + s_i) + \sum_{i \in N} \lfloor m_i \rfloor (\overline{x}_i - s_i) \\
& + g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i \in M} \lceil m_i \rceil \underline{x}_i - \sum_{i \in N} \lfloor m_i \rfloor \overline{x}_i, \text{ with } s_i \geq 0\ \forall i \\
=\ & \sum_{i \in M} \lceil m_i \rceil s_i - \sum_{i \in N} \lfloor m_i \rfloor s_i + \sum_{i=1}^n m_i \tilde{x}_i + b \\
\geq\ & \sum_{i \in M} m_i s_i - \sum_{i \in N} m_i s_i + \sum_{i=1}^n m_i \tilde{x}_i + b \\
=\ & \sum_{i \in M} m_i (\underline{x}_i + s_i) + \sum_{i \in N} m_i (\overline{x}_i - s_i) + b \\
=\ & \sum_{i=1}^n m_i x_i + b.x
\end{aligned}
$$

□

Just as the univariate class-2 estimators were shown tighter than the class-1 estimators, the multivariate class-2 estimators are tighter than their corresponding class-1 multivariate estimators. The first result shows that an optimal class-1 estimator is always less tight than its corresponding class-2 estimator (if it exists).

Theorem 11 (Relative Tightness of Class 1 and 2 Safe Overestimating Hyperplanes). *Let g be an n-variate function with linear overestimator*

$\sum_{i=1}^{n} m_i x_i + b$ over $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$. Let g be such that $g(\tilde{x}_1, \ldots, \tilde{x}_n) = \sum_{i=1}^{n} m_i \tilde{x}_i + b$ where $\tilde{x}_i \in \{\underline{x}_i, \overline{x}_i\}$. Suppose that $\sum_{i \in M} \lceil m_i \rceil x_i + \sum_{i \in N} \lfloor m_i \rfloor x_i + \lceil b + \sum_{i=1}^{n} \delta_i \rceil$ is the optimal class-1 estimator where $M = \{i | \tilde{x}_i = \underline{x}_i\}$ and $N = \{i | \tilde{x}_i = \overline{x}_i\}$. The corresponding class-2 estimator is tighter than the class-1 estimator (modulo rounding errors).

Proof. When the difference in error between the class-1 and class-2 estimators, ΔE is positive, the class-2 estimator is tighter. We calculate the difference in error of the slightly tighter estimators:

$$\begin{aligned}
\Delta E = {} & \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \Bigg\{ \sum_{i \in M} \lceil m_i \rceil x_i + \sum_{i \in N} \lfloor m_i \rfloor x_i + b + \sum_{i=1}^{n} \delta_i \\
& - \Bigg(\sum_{i \in M} \lceil m_i \rceil x_i + \sum_{i \in N} \lfloor m_i \rfloor x_i + g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i \in M} \lceil m_i \rceil \underline{x}_i - \sum_{i \in N} \lfloor m_i \rfloor \overline{x}_i \Bigg) \Bigg\} \\
= {} & \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \Bigg\{ \sum_{i=1}^{n} \delta_i - \sum_{i=1}^{n} m_i \tilde{x}_i + \sum_{i \in M} \lceil m_i \rceil \underline{x}_i + \sum_{i \in N} \lfloor m_i \rfloor \overline{x}_i \Bigg\} \\
= {} & \underbrace{\prod_{j=1}^{n} (\overline{x}_j - \underline{x}_j)}_{=P \geq 0} \Bigg\{ \sum_{i=1}^{n} \delta_i + \sum_{i \in M} (\lceil m_i \rceil - m_i) \underline{x}_i - \sum_{i \in N} (m_i - \lfloor m_i \rfloor) \overline{x}_i \Bigg\}
\end{aligned}$$

Let $I = \{i | \underline{x}_i \geq 0\}$, $J = \{i | \overline{x}_i \leq 0\}$, $K = \{i | 0 \in (\underline{x}_i, \overline{x}_i), |\underline{x}_i| \leq |\overline{x}_i|\}$ and $L = \{i | 0 \in (\underline{x}_i, \overline{x}_i), |\underline{x}_i| > |\overline{x}_i|\}$. These sets partition $\{1, \ldots, n\}$ by the rules for optimality of class-1 estimators.

$$\begin{aligned}
n\Delta E = P \Bigg\{ & \sum_{i \in K} -\mathtt{err}(m_i) \underline{x}_i + \sum_{i \in L} \mathtt{err}(m_i) \overline{x}_i \\
& + \sum_{i \in I \cup K} (\lceil m_i \rceil - m_i) \underline{x}_i - \sum_{i \in J \cup L} (m_i - \lfloor m_i \rfloor) \overline{x}_i \Bigg\}
\end{aligned}$$

Using the fact that

$$\sum_{i \in I \cup K} (\lceil m_i \rceil - m_i) \underline{x}_i \geq \sum_{i \in I} (\lceil m_i \rceil - m_i) \underline{x}_i + \sum_{i \in K} \mathtt{err}(m_i) \underline{x}_i$$

and

$$\sum_{i \in J \cup L} (m_i - \lfloor m_i \rfloor) \overline{x}_i \leq \sum_{i \in J} (m_i - \lfloor m_i \rfloor) \overline{x}_i + \sum_{i \in L} \mathtt{err}(m_i) \overline{x}_i$$

we get the following lower bound:

$$\Delta E \geq P \Bigg\{ \sum_{i \in I} (\lceil m_i \rceil - m_i) \underline{x}_i - \sum_{i \in J} (m_i - \lfloor m_i \rfloor) \overline{x}_i \Bigg\} \geq 0$$

Therefore, the class-2 estimator is tighter. □

Theorem 15 in Appendix A compares an optimal class-1 estimator with a class-2 operator that is not its direct counterpart. Once again, it provides a reasonable justification for preferring class-2 estimators in general.

6 Safe Linear Programs

There is one last issue that must be discussed when deriving a safe linear program relaxing a global optimization problem: the fact that the coefficients in the statement may be uncertain. Indeed, these coefficients may be given by intervals or they may be given by expressions and/or textual representations which are also subject to rounding errors when evaluated. As a consequence, the linear estimators described earlier generally produce a linear program whose coefficients are intervals and the postprocessing suggested in [10] does not apply directly.

This section discusses how linear constraints with interval coefficients can be safely approximated by linear constraints with floating-point coefficients. For example, it shows how to approximate a constraint of the form

$$\sum [\underline{a_i}, \overline{a_i}]\ x_i \ + \ b \leq 0.$$

by a constraint

$$\sum \tilde{a}_i\ x_i \ + \ \tilde{b} \leq 0.$$

where $\tilde{a}_i$ and $\tilde{b}$ are floating-point numbers. Interestingly, this approximation can be obtained by specializing the class-1 and class-2 estimators presented earlier. In the following, we use $\mathbb{IF}$ to denote the set of intervals whose bounds are floating-point numbers.[3]

Definition 5 (Safe Linear Programs). *Let $\mathcal{P}$ be an interval linear program*

$$\begin{array}{ll} \textit{min} & \mathbf{c}^T\mathbf{x} \\ \textit{subject to} & \\ & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}] \end{array}$$

where $\mathbf{A} \in \mathbb{IF}^{m\times n}$, $\mathbf{b} \in \mathbb{IF}^m$, and $\mathbf{c} \in \mathbb{IF}^n$. A safe linear program approximating $\mathcal{P}$ is a linear program

$$\begin{array}{ll} \textit{min} & \tilde{\mathbf{c}}^T\mathbf{x} + \tilde{d} \\ \textit{subject to} & \\ & \tilde{\mathbf{A}}\mathbf{x} \leq \tilde{\mathbf{b}} \\ & \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}] \end{array}$$

[3] Note that the estimators presented in this section naturally generalize to unbounded intervals. However, we omit these here since the postprocessing step [10] requires bounded intervals to correct the approximate solution given by commercial codes.

where $\tilde{\mathbf{A}} \in \mathbb{F}^{m\times n}$, $\tilde{\mathbf{b}} \in \mathbb{F}^m$, $\tilde{\mathbf{c}} \in \mathbb{F}^m$, *and* $\tilde{d} \in \mathbb{F}$ *satisfying*

$$\{\mathbf{x}|\mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}]\} \subseteq \{\mathbf{x}|\tilde{\mathbf{A}}\mathbf{x} \leq \tilde{\mathbf{b}}, \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}]\}$$

and

$$\forall \mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}] : \; \tilde{\mathbf{c}}^T\mathbf{x} + \tilde{d} \leq \mathbf{c}^T\mathbf{x}$$

We now show that a safe linear program can be obtained without increasing the number of variables and constraints. Since the program is linear, it is sufficient to consider each variable's interval coefficient independently. Figure 7 illustrates the intuition underlying the most complicated case.

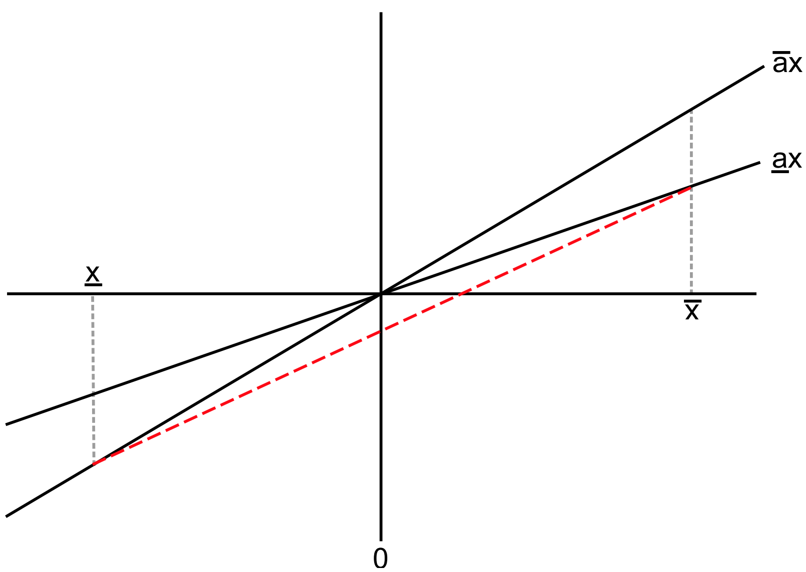

Fig. 7. Finding a safe linear representation.

Theorem 12 (Safe Linear Underestimators for Interval Linear Terms). *Let* $[\underline{a}, \overline{a}]x$ *be a term appearing within an interval linear program and* $x \in [\underline{x}, \overline{x}]$. *We have*

$$\forall a \in [\underline{a}, \overline{a}] : ax \geq \begin{cases} \underline{a}x & \underline{x} \geq 0 \\ \overline{a}x & \overline{x} \leq 0 \\ \underline{a}x + \lfloor \mathtt{err}(a)\underline{x} \rfloor & \underline{x} < 0 \\ \overline{a}x - \lfloor \mathtt{err}(a)\overline{x} \rfloor & \overline{x} > 0 \\ \underline{mx} + \lfloor \overline{a}\underline{x} - \underline{mx} \rfloor & |\underline{x}| < |\overline{x}|, \;\; 0 \in [\underline{x}, \overline{x}] \\ \overline{mx} + \lfloor \underline{a}\overline{x} - \overline{mx} \rfloor & |\underline{x}| \geq |\overline{x}|, \;\; 0 \in [\underline{x}, \overline{x}] \end{cases}$$

where

$$\underline{m} = \max\left(\underline{a}, \lfloor \frac{\underline{a}\overline{x} - \overline{a}\underline{x}}{\overline{x} - \underline{x}} \rfloor\right) \ \text{and}\ \overline{m} = \min\left(\overline{a}, \lceil \frac{\underline{a}\overline{x} - \overline{a}\underline{x}}{\overline{x} - \underline{x}} \rceil\right).$$

Proof. The first four cases are safe: they are class 1 estimators with $\underline{m} = \underline{a}$, $\overline{m} = \overline{a}$ and $b = 0$. The last two cases are safe: they are class 2 estimators with $m = \frac{\underline{a}\overline{x} - \overline{a}\underline{x}}{\overline{x} - \underline{x}}$ (the slope of the line joining $(\underline{x}, \overline{a}\underline{x})$ and $(\overline{x}, \underline{a}\overline{x})$. The selection rule for the last two cases comes from the theorems on class 2 estimators. □

Theorem 13 (Tightness of Linear Underestimators for Interval Terms). *Theorem 12 provides the tightest possible safe linear program without the addition of new variables.*

Proof. The first two cases are trivially the tightest possible. Consider the tightness of the final two cases. The safe estimating line depicted in Figure 7 is the tightest possible. When $\lfloor \frac{\underline{a}\overline{x} - \overline{a}\underline{x}}{\overline{x} - \underline{x}} \rfloor > \underline{a}$ (resp. $\lceil \frac{\underline{a}\overline{x} - \overline{a}\underline{x}}{\overline{x} - \underline{x}} \rceil < \overline{a}$) the fifth (resp. sixth) case is necessarily tighter than the third (fourth). When $\underline{m} = \underline{a}$ (resp. $\overline{m} = \overline{a}$), the fifth (sixth) case reduces to the third (fourth) case. □

A safe linear program can thus be obtained easily by safely approximating each interval linear term independently in the constraints and the objective function. Of course, the constants generated during this process must be collected and rounded down as added in order to maintain safety.

7 Conclusion

Global optimization problems are often approached by branch and bound algorithms which use linear relaxations of the nonlinear constraints computed from the current variable bounds at each node of the search tree. This paper considered the problem of obtaining safe linear relaxations which are guaranteed to enclose the solutions of the nonlinear problem. It contains three main contributions. First, it studied two classes of linear estimators for univariate functions. The first class of estimators generalizes the results in [9] from quadratic to arbitrary functions, while the second class is entirely novel. Theoretical and numerical results indicated that class-2 estimators, when they apply, are tighter than class-1 estimators. Second, the paper generalized the univariate results to multivariate functions and indicated how to derive estimators for multivariate functions by combining univariate estimators derived for each variable independently. Moreover, it showed that the combination of tight class-1 safe univariate estimators is a tight class-1 safe multivariate estimators and class-2 safe multivariate estimators are tighter than their corresponding optimal class-1 safe multivariate estimators. Finally, the paper showed how to safely approximate linear programs with interval coefficients.

In conjunction with the safe bounds on the linear relaxations derived in [10, 4], these results provide a comprehensive framework to derive safe linear estimators for global optimization problems, laying the theoretical foundation for rigorous results of branch and bound approaches to global optimization based on linear programming.

Acknowledgements

This work is partially supported by NSF ITR Awards DMI-0121495 and ACI-0121497 and by a Canadian NSERC PGS-A grant.

References

1. F. Benhamou, D. McAllester, and P. Van Hentenryck. CLP(Intervals) Revisited. In *Proceedings of the International Symposium on Logic Programming (ILPS-94)*, pages 124–138, Ithaca, NY, November 1994.
2. F. Benhamou and W. Older. Applying Interval Arithmetic to Real, Integer and Boolean Constraints. *Journal of Logic Programming*, 1997. To appear.
3. I. E. Grossmann and S. Lee. Generalized convex disjunctive programming: Nonlinear convex hull relaxation. Submitted for publication (2001).
4. C. Jansson. Rigorous lower and upper bounds in linear programming, 2002.
5. Y. Lebbah, M. Rueher, and C. Michel. A global filtering algorithm for handling systems of quadratic equations and inequations. *Lecture Notes in Computer Science*, 2470:109–123, 2002.
6. S. Lee and I. E. Grossmann. A global optimization algorithm for nonconvex generalized disjunctive programming and applications to process systems. *Computers and Chemical Engineering*, 25:1675–1697, 2001.
7. O. Lhomme. Consistency Techniques for Numerical Constraint Satisfaction Problems. In *Proceedings of the 1993 International Joint Conference on Artificial Intelligence*, Chamberry, France, August 1993.
8. G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I - convex underestimating problems. *Mathematical Programming*, 10:146–175, 1976.
9. C. Michel, Y. Lebbah, and M. Rueher. Safe embedding of the simplex algorithm in a CSP framework. *Proceedings CPAIOR'03*, pages 210–220, 2003.
10. A. Neumaier and O. Shcherbina. Safe bounds in linear and mixed-integer programming. to appear in Mathematical Programming.
11. I. Quesada and I. E. Grossmann. A global optimization algorithm for linear fractional and bilinear programs. *Journal of Global Optimization*, 6:39–76, 1995.
12. H. S. Ryoo and N. V. Sahinidis. Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Computer and Chemical Engineering*, 19:551–566, 1995.
13. P. Van Hentenryck, D. McAllister, and D. Kapur. Solving Polynomial Systems Using a Branch and Prune Approach. *SIAM Journal on Numerical Analysis*, 34(2), 1997.
14. P. Van Hentenryck, L. Michel, and Y. Deville. *Numerica: a Modeling Language for Global Optimization*. The MIT Press, Cambridge, Mass., 1997.
15. J. M. Zamora and I. E. Grossmann. A comprehensive global optimization approach for the synthesis of heat exchanger networks with no stream splits. *Computers and Chemical Engineering*, 21:S65–S70, 1997.
16. J. M. Zamora and I. E. Grossmann. A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. *Journal of Global Optimization*, 14:217–249, 1999.

A Additional Tightness Results

For completeness, we compare the class-2 safe linear overestimator $\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil$ with the class-1 estimators that uses $m^* = \lfloor m \rfloor$. This result is not useful in practice when the class-2 operator using $\lfloor m \rfloor$ is available (which is always the case for secant lines for instance). In this case, one should choose the other class-2 estimator which is guaranteed to be tighter theoretically. However, the result sheds some light on the relationships between the two classes of estimators and indicate that, in general, class-2 estimators should be preferred.

Theorem 14 (Relative Tightness of Class 1 and Class 2 Safe Linear Overestimators). *Let g be a univariate function with linear overestimator $mx + b$ over $[\underline{x}, \overline{x}]$ such that $g(\underline{x}) = m\underline{x} + b$. The class-2 estimator $(\lceil m \rceil x + \lceil g(\underline{x}) - \lceil m \rceil \underline{x} \rceil)$ is often tighter than the optimal (using the rules given in Theorems 3 and 4) class-1 estimator when $|\overline{x}| \leq |\underline{x}|$ (modulo rounding errors).*

Proof. Again we prove the result modulo rounding errors. Consider the difference in error (given by definition 3) ΔE between the class-1 estimator $\lfloor m \rfloor x + \lceil b + \mathtt{err}(m)\overline{x} \rceil$ when $|\underline{x}| \geq |\overline{x}|$ and $0 \in (\underline{x}, \overline{x})$ (the conditions for optimality of this estimator) and the class-2 estimator. The class-2 estimator is tighter when $\Delta E > 0$. We consider the difference in error between the slightly tighter estimators:

$$
\begin{aligned}
\Delta E &= \int_{\underline{x}}^{\overline{x}} \lfloor m \rfloor x + b + \mathtt{err}(m)\overline{x} - (\lceil m \rceil x + g(\underline{x}) - \lceil m \rceil \underline{x}) \mathrm{d}x \\
&= -\frac{1}{2}\mathtt{err}(m)(\overline{x}^2 - \underline{x}^2) + (\mathtt{err}(m)\overline{x} + (\lceil m \rceil - m)\underline{x})(\overline{x} - \underline{x}) \\
&= (\overline{x} - \underline{x}) \left\{ \frac{1}{2}\mathtt{err}(m)(\overline{x} - \underline{x}) + (\lceil m \rceil - m)\underline{x} \right\}
\end{aligned}
$$

The above is positive when

$$
\frac{1}{2}\mathtt{err}(m)(\overline{x} - \underline{x}) + (\lceil m \rceil - m)\underline{x} \geq 0
$$

or, alternatively, when

$$
\frac{\lceil m \rceil - m}{\mathtt{err}(m)} \leq \frac{|\overline{x}| + |\underline{x}|}{2|\underline{x}|} \in [0.5, 1],
$$

where the final range is given by the range for $\overline{x}$: $[0, |\underline{x}|]$.

Assuming that m is uniformly distributed in $[\lfloor m \rfloor, \lceil m \rceil]$, the class-2 estimator is expected to be tighter in at least 50% of the cases and the percentage tends to 100% as $\overline{x} \to |\underline{x}|$. □

The result generalizes to the multivariate case and suggests that, in general, class-2 estimators should be preferred.

Theorem 15 (Relative Tightness of Class 1 and 2 Safe Overestimating Hyperplanes). *Let g be an n-variate function with linear overestimator $\sum_{i=1}^{n} m_i x_i + b$ over $[\underline{x}_1, \overline{x}_1] \times \cdots \times [\underline{x}_n, \overline{x}_n]$ such that $g(\tilde{x}_1, \ldots, \tilde{x}_n) = \sum_{i=1}^{n} m_i \tilde{x}_i + b$ where $\tilde{x}_i \in \{\underline{x}_i, \overline{x}_i\}$. The corresponding class-2 estimator is often tighter than the optimal class-1 estimator.*

Proof. The optimal class-1 estimator is:

$$\sum_{i=1}^{n} m_i^* x_i + \lceil b + \sum_{i=1}^{n} \delta_i^* \rceil, \ m_i^* \in \{\lceil m \rceil, \lfloor m \rfloor\}$$

The class-2 estimator in consideration is:

$$\sum_{i=1}^{n} \tilde{m}_i x_i + \lceil g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i=1}^{n} \tilde{m}_i \tilde{x}_i \rceil, \ \tilde{m}_i = \begin{cases} \lceil m \rceil \text{ if } \tilde{x}_i = \underline{x}_i, \\ \lfloor m \rfloor \text{ if } \tilde{x}_i = \overline{x}_i. \end{cases}$$

Consider the difference in error, ΔE, between the slightly tighter estimators:

$$\begin{aligned}
\Delta E &= \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \quad \Bigg\{ \sum_{i=1}^{n} m_i^* x_i + b + \sum_{i=1}^{n} \delta_i^* \\
&\qquad\qquad - \left(\sum_{i=1}^{n} \tilde{m}_i x_i + g(\tilde{x}_1, \ldots, \tilde{x}_n) - \sum_{i=1}^{n} \tilde{m}_i \tilde{x}_i \right) \Bigg\} \\
&= \int_{\underline{x}_1}^{\overline{x}_1} \mathrm{d}x_1 \cdots \int_{\underline{x}_n}^{\overline{x}_n} \mathrm{d}x_n \quad \left\{ \sum_{i=1}^{n} (m_i^* - \tilde{m}_i) x_i + \delta_i^* - (m_i - \tilde{m}_i) \tilde{x}_i \right\} \\
&= \underbrace{\prod_{j=1}^{n} (\overline{x}_j - \underline{x}_j)}_{=P \geq 0} \sum_{i=1}^{n} \frac{1}{2} (m_i^* - \tilde{m}_i)(\overline{x}_i + \underline{x}_i) + \delta_i^* - (m_i - \tilde{m}_i) \tilde{x}_i
\end{aligned}$$

The class-2 estimator is tighter when $\Delta E \geq 0$. First define the following sets which partition $\{1, \ldots, n\}$ for the optimal class-1 estimator: $I = \{i | \underline{x}_i \geq 0\}$, $J = \{i | \overline{x}_i \leq 0\}$, $K = \{i | 0 \in (\underline{x}_i, \overline{x}_i), |\underline{x}_i| \leq |\overline{x}_i|\}$ and $L = \{i | 0 \in (\underline{x}_i, \overline{x}_i), |\underline{x}_i| > |\overline{x}_i|\}$. Also define a partition for the class-2 estimators: $M = \{i | \tilde{x}_i = \overline{x}_i\}$ and

$N = \{i|\tilde{x}_i = \underline{x}_i\}$. The difference in error becomes:

$$\begin{aligned}
\Delta E = P \Bigg\{ & \sum_{i\in(I\cup K)\cap M} \frac{1}{2}\mathtt{err}(m_i)(\overline{x}_i + \underline{x}_i) - \sum_{i\in(J\cup L)\cap N} \frac{1}{2}\mathtt{err}(m_i)(\overline{x}_i + \underline{x}_i) \\
& - \sum_{i\in K} \mathtt{err}(m_i)\underline{x}_i + \sum_{i\in L} \mathtt{err}(m_i)\overline{x}_i \\
& - \sum_{i\in M} (m_i - \lfloor m_i \rfloor)\overline{x}_i + \sum_{i\in N} (\lceil m_i \rceil - m_i)\underline{x}_i \Bigg\} \\
= P \Bigg\{ & \sum_{i\in((I\cup K)\cap M)\cup((J\cup L)\cap N)} \frac{1}{2}\mathtt{err}(m_i)(|\overline{x}_i| + |\underline{x}_i|) \\
& - \sum_{i\in K\cap N} \mathtt{err}(m_i)\underline{x}_i + \sum_{i\in L\cap M} \mathtt{err}(m_i)\overline{x}_i \\
& - \sum_{i\in M} (m_i - \lfloor m_i \rfloor)\overline{x}_i + \sum_{i\in N} (\lceil m_i \rceil - m_i)\underline{x}_i \Bigg\}
\end{aligned}$$

We now use the assumptions that:

$$\begin{aligned}
\frac{1}{2}\mathtt{err}(m_i)(|\overline{x}_i| + |\underline{x}_i|) \geq |\overline{x}_i|(m_i - \lfloor m_i \rfloor),\ i \in M \\
\frac{1}{2}\mathtt{err}(m_i)(|\overline{x}_i| + |\underline{x}_i|) \geq |\underline{x}_i|(\lceil m_i \rceil - m_i),\ i \in N
\end{aligned}$$

These assumptions appear in the proof of Theorem 14 and the corresponding theorem for $g(\overline{x}) = m\overline{x} + b$. The assumptions, as argued in Theorem 14, hold more often as $|\overline{x}_i| \to |\underline{x}_i|$ when $i \in N$ and as $|\underline{x}_i| \to |\overline{x}_i|$ when $i \in M$. Using these assumptions, we find that:

$$\begin{aligned}
\Delta E \geq P \Bigg\{ & \sum_{i\in L\cap M} \mathtt{err}(m_i)\overline{x}_i - \sum_{i\in K\cap N} \mathtt{err}(m_i)\underline{x}_i \\
& - \sum_{i\in(J\cup L)\cap M} (m_i - \lfloor m_i \rfloor)\overline{x}_i + \sum_{i\in(I\cup K)\cap N} (\lceil m_i \rceil - m_i)\underline{x}_i \Bigg\} \\
= P \Bigg\{ & \sum_{i\in L\cap M} (\lceil m_i \rceil - m_i)\overline{x}_i + \sum_{i\in J\cap M} (\lfloor m_i \rfloor - m_i)\overline{x}_i \\
& + \sum_{i\in K\cap N} (\lfloor m_i \rfloor - m_i)\underline{x}_i + \sum_{i\in I\cap N} (\lceil m_i \rceil - m_i)\underline{x}_i \Bigg\}
\end{aligned}$$

Upon further analysis, we find that $\Delta E \geq 0$. The class-2 estimator is tighter than the optimal class-1 estimator when the above assumptions hold. Since these conditions are biased to be met more often than not, the class-2 estimator is often the tightest estimator. □