Supplement to "A Framework For Estimation of Convex Functions"

T. Tony Cai* and Mark G. Low
Department of Statistics
The Wharton School
University of Pennsylvania

Abstract

In this supplement we prove the additional technical lemmas stated in Section 7.1 which are used in the proofs of the main results.

Lemma 4 *The function H^{-1} defined in Section 2.1 is concave and nondecreasing. It is strictly increasing for all x where $H^{-1}(x) < \frac{1}{2}$. Moreover for $C \geq 1$ it satisfies*

$$H^{-1}(Ct) \leq C^{\frac{2}{3}} H^{-1}(t). \tag{60}$$

The function K defined in Section 2.1 is also increasing and satisfies for $C \geq 1$

$$C^{\frac{2}{3}} K(t) \leq K(Ct) \leq CK(t). \tag{61}$$

Proof of Lemma 4: First note that H is a nondecreasing convex function. Moreover there is a unique point x_0 such that it is strictly increasing on some open interval $(x_0, \frac{1}{2})$ where $f_s(x_0) = 0$. The inverse function $H^{-1}(x)$ is thus strictly increasing on the interval $(0, H(\frac{1}{2}))$. In this interval $H^{-1}(x) < \frac{1}{2}$. For $x > H(\frac{1}{2})$, $H^{-1}(x) = \frac{1}{2}$. It follows that H^{-1} is nondecreasing. The concavity of H^{-1} is guaranteed because it is the inverse of an increasing convex function.

Now let $C \geq 1$. Then since f_s is convex and $f_s(0) = 0$ it follows that whenever $C^{2/3} y \leq \frac{1}{2}$,

$$C^{2/3} f_s(y) \leq f_s(C^{2/3} y)$$

and hence also

$$CH(y) = C\sqrt{y} f_s(y) \leq C^{1/3} \sqrt{y} f_s(C^{2/3} y) = H(C^{2/3} y).$$

*The research of Tony Cai was supported in part by NSF Grant DMS-0604954 and NSF FRG Grant DMS-0854973.

Now let $y = H^{-1}(t)$. Clearly if $C^{2/3}H^{-1}(t) \geq \frac{1}{2}$ then (60) must hold. Hence suppose that $C^{2/3}H^{-1}(t) < \frac{1}{2}$. In this case let $y = H^{-1}(t)$ Then

$$CH(H^{-1}(t)) \leq H(C^{2/3}H^{-1}(t))$$

and hence

$$Ct \leq H(C^{2/3}H^{-1}(t)).$$

Consequently,

$$H^{-1}(Ct) \leq H^{-1}(H(C^{2/3}H^{-1}(t))) = C^{2/3}H^{-1}(t)$$

which establishes (60) in this other case.

Note that for $C \geq 1$,

$$K(Ct) = \frac{Ct}{\sqrt{H^{-1}(Ct)}} \geq \frac{Ct}{C^{1/3}\sqrt{H^{-1}(t)}}.$$

The first inequality in equation (61) and the fact that K is increasing immediately follows. On the other hand,

$$K(Ct) = \frac{Ct}{\sqrt{H^{-1}(Ct)}} \leq \frac{Ct}{\sqrt{H^{-1}(t)}} = CK(t),$$

which yields the second inequality in equation (61). ∎

Lemma 5 *Let f be a nonnegative convex function on $[-\frac{1}{2}, \frac{1}{2}]$. For $d > 0$ let t be the supremum over all y with $f_s(y) \leq d$ where f_s defined in Section 2.1 is the symmetrized and centered version of f. Then there is a convex function g with $g(0) - f(0) = d$ and for which*

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (g(x) - f(x))^2 dx \leq \frac{9}{4} d^2 t. \tag{62}$$

It follows that for each $0 \leq t \leq \frac{1}{2}$ there is a convex function g with $g(0) - f(0) = f_s(t)$ such that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (g(x) - f(x))^2 dx \leq \frac{9}{4} H^2(t) \tag{63}$$

where the function H is defined in Section 2.1 Moreover for any convex h with $h(0) - f(0) = d > 0$

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (h(x) - f(x))^2 dx \geq \frac{2}{3} d^2 t. \tag{64}$$

Remark: The constants $\frac{9}{4}$ and $\frac{2}{3}$ in (62) and (64) are sharp.

Proof of Lemma 5: Throughout this proof we shall without loss of generality take $f(0) = 0$. First suppose that $f_s(\frac{1}{2}) < d$. Then $t = \frac{1}{2}$. In this case take $g(x) = d$ and it is clear that

$$\int_{-1/2}^{\frac{1}{2}} (g(x) - f(x))^2 dx \leq 2d^2$$

and in this case (62) holds.

We must now consider the situation where $f_s(t) = d$ and hence $f(t) + f(-t) = 2d$. We shall consider two cases. In the first $\max(f(t), f(-t)) \geq \frac{3d}{2}$ and in the second case $d \leq \max(f(t), f(-t)) < \frac{3d}{2}$. In the first case for the moment assume that $f(t) \geq \frac{3d}{2}$. Then take $g(x) = \max(f(x), d + \frac{d}{2t}x)$. Note that g is convex as it is a maximum of two convex functions. Also $g(x) = f(x)$ at least for $x \leq -2t$ and $x \geq t$. Moreover since f is nonnegative It is also clear that $g(x) - f(x) \leq d + \frac{d}{2t}x$. Hence in this case

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (g(x) - f(x))^2 dx \leq \int_{-2t}^{t} (d + \frac{d}{2t}x)^2 dx = \frac{9}{4}d^2 t.$$

Similarly when $f(-t) \geq \frac{3d}{2}$ an entirely similar argument can be applied to the function $g(x) = \max(f(x), d - \frac{d}{2t}x)$. Equation (62) of the lemma thus holds under the first case.

In the second case we have $\max(f(t), f(-t)) \leq \frac{3d}{2}$. In this case take $g(x) = \max(f(x), d + \frac{f(t)-f(-t)}{2t}x)$. In this case $g(x) = f(x)$ for $|x| \geq t$ and otherwise $g(x) - f(x) \leq d + \frac{f(t)-f(-t)}{2t}x$. It follows that

$$\begin{aligned}
\int_{-\frac{1}{2}}^{\frac{1}{2}} (g(x) - f(x))^2 dx &\leq \int_{-t}^{t} (d + \frac{f(t) - f(-t)}{2t}x)^2 dx \\
&= \frac{2t}{3}\left(3d^2 + \frac{(f(t) - f(-t))^2}{4}\right) \\
&\leq \frac{2t}{3}\left(3d^2 + \frac{d^2}{4}\right) = \frac{13}{6}d^2 t.
\end{aligned}$$

Thus equation (62) of the lemma also holds in the second case since $\frac{13}{6} \leq \frac{9}{4}$. Equation (63) follows immediately on taking $d = f(t)$ and noting that $t^2 f(t) = H(t)$. We now turn to the proof of (64). For any pair of convex functions f and h let $\tilde{f}(x) = \frac{f(x)+f(-x)}{2}$ and $\tilde{h}(x) = \frac{h(x)+h(-x)}{2}$ be symmetrized versions. Note that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (h(x) - f(x))^2 dx \geq \int_{-\frac{1}{2}}^{\frac{1}{2}} (\tilde{h}(x) - \tilde{f}(x))^2 dx.$$

Note that since $\tilde{h}$ is convex and symmetric with $\tilde{h}(0) = d$ it follows that $\tilde{h}(x) \geq d$ for all $x \in [-1/2, 1/2]$. Hence $\tilde{h}(x) \geq d$. Note also that $f_s(t) \leq d$ and hence for $|x| \leq t$ it follows that $\tilde{f}(x) \leq \frac{|x|}{t}d$ and hence

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} (\tilde{h}(x) - \tilde{f}(x))^2 dx \geq \int_{-t}^{t} (d - \frac{|x|}{t}d)^2 dx = \frac{2}{3}d^2 t$$

and (64) also follows. ∎

Lemma 6 *Let f and g be convex functions with $f(0)-g(0)=a>0$. Let t be the supremum of all y for which $f_s(y)\le a$. Then*

$$\int_{-1/2}^{1/2}(f(x)-g(x))^2dx \ge 0.3ta^2. \tag{65}$$

The proof of this lemma requires an additional technical result which will be stated and proved in Lemma 11 at the end of this supplement.

Proof of Lemma 6: Note that, since as in the proof of lemma 5,

$$\int_{-1/2}^{1/2}(\tilde f(x)-\tilde g(x))^2dx \le \int_{-1/2}^{1/2}(f(x)-g(x))^2dx$$

where $\tilde f(x)=\frac{f(x)+f(-x)}{2}$ and $\tilde g(x)=\frac{g(x)+g(-x)}{2}$, it suffices to prove the lemma for all symmetric convex functions f and g. Hence we shall assume f and g to be convex, even functions and without loss of generality we shall also take $f(0)=0$ and hence $g(0)=-a$ and $f(t)\le a$. First suppose that $g(x)\le 0$ for $0\le x\le \frac{t}{2}$. Then for $0\le x\le\frac{t}{2}$, $g(x)\le\frac{2a}{t}(x-\frac{t}{2})$. In this case

$$\int_{-1/2}^{1/2}(f(x)-g(x))^2dx \ge \int_{-t/2}^{t/2}\frac{4a^2}{t^2}(x-\frac{t}{2})^2dx = \frac{1}{3}ta^2$$

and the Lemma would hold in this case. So suppose that $g(t_1)=0$ where $t_1<\frac{t}{2}$. In this case f and g must meet in one and only one point. Suppose that $f(t_2)=g(t_2)$. Now let $h(x)=-a+\frac{f(t_2)+a}{t_2}x$. Note that $g(x)\le h(x)\le f(x)$ for $0\le x\le t_2$ and that $g(x)\ge h(x)\ge f(x)$ for $t_2\le x\le\frac{1}{2}$. It follows that

$$\int_{-1/2}^{1/2}(f(x)-g(x))^2dx = 2\int_0^{1/2}(f(x)-g(x))^2dx \ge 2\int_0^t(f(x)-h(x))^2dx.$$

Now let

$$k(x)=\max\left\{\frac{f(t)-f(t_2)}{t-t_2}(x-t)+f(t),\ 0\right\}.$$

Since $f(x)\ge k(x)\ge h(x)$ for $0\le x\le t_2$ and $h(x)\ge k(x)\ge f(x)$ for $t_2\le t$ it follows that

$$\int_{-1/2}^{1/2}(f(x)-g(x))^2dx \ge 2\int_0^t(k(x)-h(x))^2dx$$

Now let $y=\frac{x}{t}$. Note that $k(0)=0$ and $h(0)=-a$ and that k and h are of the form of the functions in Lemma 11. It then follows from this lemma that

$$2\int_0^t(k(x)-h(x))^2dx = 2\int_0^1(k(ty)-h(ty))^2tdy \ge 2*0.1572ta^2 \ge 0.3ta^2 \quad ∎$$

Lemma 7 *Set* $\sigma_{j_*} = \frac{2^{(j_*-1)/2}}{\sqrt{n}}$. *Let* j_* *be defined as in Section 4, then*

$$ET_{j_*} \le \min(E\delta_{j_*} - f(0), \sigma_{j_*}). \tag{66}$$

For $k \ge 1$,

$$E\delta_{j_*-k} - f(0) \ge 2^{k-\frac{3}{2}}\sigma_{j_*} \tag{67}$$

and

$$ET_{j_*-k} \ge \frac{2^{k-1}}{\sqrt{6}}\sigma_{j_*}. \tag{68}$$

Proof of Lemma 7: The proof of this lemma will partly use Lemma 4. First note that Lemma 4 gives $ET_{j_*} \le E\delta_{j_*} - f(0)$ and so (66) is clear in the case that $E\delta_{j_*} - f(0) \le \sigma_{j_*}$. In the case $E\delta_{j_*} - f(0) = \lambda\sigma_{j_*}$ where $\lambda > 1$ note that since δ_{j_*} has the smallest mean squared error it follows that

$$E\delta_{j_*+1} - f(0) \ge \sqrt{(\lambda^2 - 1)}\sigma_{j_*}$$

and hence

$$ET_{j_*} = E\delta_{j_*} - E\delta_{j_*+1} \le (\lambda - \sqrt{(\lambda^2 - 1)})\sigma_{j_*}.$$

This last expression is a decreasing function in λ when $\lambda \ge 1$ and so

$$ET_{j_*} \le \sigma_{j_*}$$

showing (66) in this other case. Now suppose that

$$E\delta_{j_*} - f(0) = \lambda\sigma_{j_*}.$$

Then

$$(E\delta_{j_*-1} - f(0))^2 + \frac{2^{j_*-2}}{n} \ge (E\delta_{j_*} - f(0))^2 + \frac{2^{j_*-1}}{n}$$

and hence

$$E\delta_{j_*-1} - f(0) \ge \sqrt{\frac{1}{2} + \lambda^2}\sigma_{j_*}$$

and

$$ET_{j_*-1} \ge (\sqrt{\frac{1}{2} + \lambda^2} - \lambda)\sigma_{j_*}.$$

Then equation (67) immediately follows from Lemma 4. Note also that Lemma 4 also yields $ET_{j_*-1} \ge \lambda\sigma_{j_*}$. Now the function

$$h(\lambda) = (\sqrt{\frac{1}{2} + \lambda^2} - \lambda).$$

is decreasing in λ and so the maximum of $h(\lambda)$ and λ occurs when $h(\lambda) = \lambda$ which has a solution of $\lambda = \frac{1}{\sqrt{6}}$ It then follows that $ET_{j_*-1} \ge \frac{1}{\sqrt{6}}$ and it follows from Lemma 4 that for $k \ge 1$ $ET_{j_*-1} \ge 2^{k-1}\frac{1}{\sqrt{6}}$ yielding equation (68). ∎

Lemma 8 *For $b > 0$ let t_b be the supremum over all t where $f_s(t) \leq br_n^{\frac{1}{2}}(f)$. Then*

$$t_b \leq \frac{2}{4-b^2}\frac{1}{nr_n(f)} \tag{69}$$

and for $b \geq \frac{2}{\sqrt{3}}$,

$$t_b \leq \frac{b3\sqrt{3}}{8nr_n(f)}. \tag{70}$$

Proof of Lemma 8: For $b > 0$ let t_b be the supremum of all points t where $f_s(t) \leq br_n^{\frac{1}{2}}(f)$. Note that

$$\frac{1}{2nt_b} + \frac{b^2 r_n(f)}{4} \geq \frac{1}{2nt_b} + (\frac{1}{t_b}\int_0^{t_b} f_s(t)dt)^2 \geq r_n(f).$$

Hence

$$t_b \leq \frac{2}{4-b^2}\frac{1}{nr_n(f)}$$

establishing (69).

Now for $b \geq \frac{2}{\sqrt{3}}$ the convexity of f_s as well as the fact that $f_s(0) = 0$ also gives the bound $t_b \leq \frac{\sqrt{3}b}{2} t_{\frac{2}{\sqrt{3}}}$ and subsituting the bound from (69) for $t_{\frac{2}{\sqrt{3}}}$ then yields (70). ∎

Lemma 9 *Let $\lambda = \sqrt{2}$ and let $h(x)$ be the function given by*

$$\begin{aligned} h(x) &= P(Z \leq \lambda - \frac{x}{2}) + \frac{0.649}{1+x^2} + \frac{1}{4}\frac{x^2}{1+x^2} \\ &+ \frac{1}{1+x^2}\sum_{m=1}^{\infty}(2^m\sqrt{3} + 2^{-m/2}2x)\left(P(Z \leq \lambda)\prod_{l=0}^{m-1} P(Z > \lambda - 2^{-3l/2}\min(x,1))\right)^{1/2}. \end{aligned}$$

Then

$$\sup_{0 \leq x \leq \frac{2}{\sqrt{3}}} h(x) \leq 4.7.$$

Proof of Lemma 9: Note that $h(x)$ is a univariate continuous function. This bound can be verified through direct numerical calculations. ∎

Lemma 10 *Let $g_m(x,y) = (x^2 + 2^{-m})P(Z \leq \lambda - 2^{m/2}(x-y))$. Then for $m \geq 2$ and $y \geq 2^{m-3/2}$*

$$\sup_{x \geq 2y} g_m(x,y) = (4y^2 + 2^{-m})P(Z \leq \lambda - 2^{m/2}y) \tag{71}$$

$$\sup_{x \geq 2y, y \geq 2^{m-3/2}} g_m(x,y) = (2^{2m-1} + 2^{-m})P(Z \geq 2^{3(m-1)/2} - \lambda). \tag{72}$$

Moreover

$$\sup_{x\geq 2y, y\geq\sqrt{2}} g_2(x,y) \leq 0.649 \tag{73}$$

$$\sup_{x\geq \max(\frac{1}{\sqrt{2}},2y), y\geq 0} \frac{g_1(x,y)}{1+y^2} \leq 1.2. \tag{74}$$

Proof of Lemma 10: For fixed $m \geq 2$ and $y \geq 2^{m-3/2}$, write $g_m(x,y)$ as a function of x,

$$\begin{aligned} g_m(x,y) &= x^2 P(Z \leq (\lambda + 2^{m/2}y) - 2^{m/2}x) + 2^{-m} P(Z \leq (\lambda + 2^{m/2}y) - 2^{m/2}x) \\ &= h_1(x) + h_2(x). \end{aligned}$$

The second term $h_2(x)$ is clearly decreasing in x and hence for $x \geq 2y$,

$$\sup_{x\geq 2y} h_2(x) = 2^{-m} P(Z \leq \lambda - 2^{m/2}y).$$

Now let us consider $h_1(x)$. Set $\tau = 2^{m/2}$ and $\gamma = \lambda + 2^{m/2}y$. Then $h_1(x) = x^2 P(Z > \tau x - \gamma)$. Then

$$h_1'(x) = 2xP(Z > \tau x - \gamma) - \tau x^2 \phi(\tau x - \gamma).$$

Hence $g'(x) \leq 0$ if

$$P(Z > \tau x - \gamma) \leq \frac{\tau x}{2}\phi(\tau x - \gamma).$$

It follows from the fact $P(Z > z) \leq z^{-1}\phi(z)$ for $z > 0$ that $h_1'(x) < 0$ if

$$\tau x(\tau x - \gamma) \geq 2.$$

This holds for

$$x \geq \frac{\gamma\tau + \sqrt{\gamma^2\tau^2 + 8\tau^2}}{2\tau^2} = \frac{\gamma + \sqrt{\gamma^2+8}}{2\tau}. \tag{75}$$

We only need to verify $2y \geq \frac{\gamma+\sqrt{\gamma^2+8}}{2\tau}$ or equivalently

$$4\tau y \geq \gamma + \sqrt{\gamma^2+8}. \tag{76}$$

Write $z = 2^{m/2}y$. Then (76) is equivalent to

$$4z \geq (\sqrt{2}+z) + \sqrt{(\sqrt{2}+z)^2 + 8}$$

which is the same as $z^2 - \sqrt{2}z - 1 \geq 0$. This last inequality holds for all $z \geq (\sqrt{2}+\sqrt{6})/2$. Note that $m \geq 2$ and so

$$z = 2^{m/2}y \geq 2^{3m/2-3/2} \geq 2\sqrt{2} \geq (\sqrt{2}+\sqrt{6})/2.$$

This proves (71).

We have shown that $h_1(x) = x^2 P(Z > \tau x - \gamma)$ is decreasing for $x \geq \frac{\gamma + \sqrt{\gamma^2+8}}{2\tau}$. It then follows easily that $(4y^2 + 2^{-m})P(Z \leq \lambda - 2^{m/2}y)$ is decreasing in y for $y \geq 2^{m-3/2}$ and so

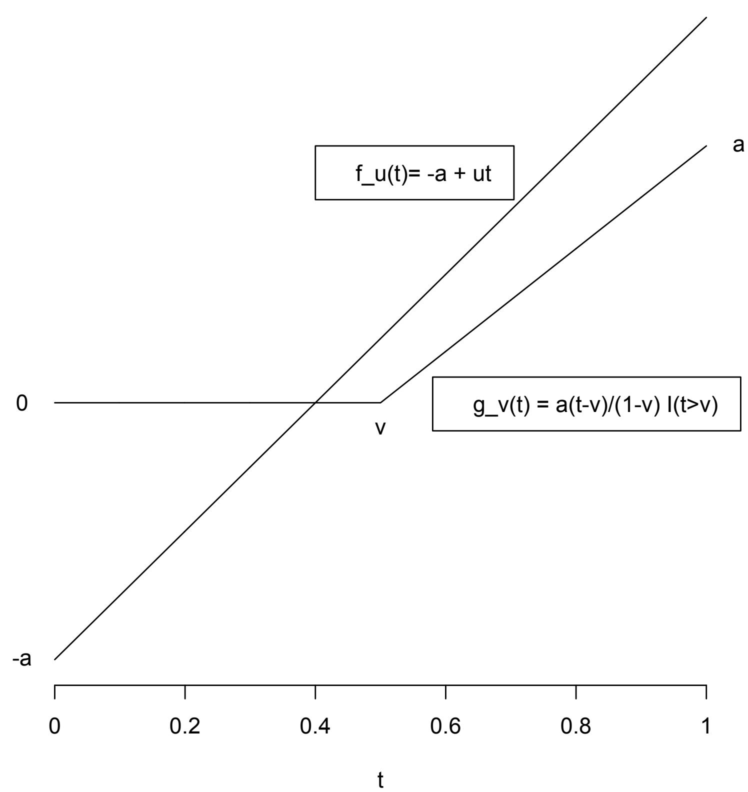
$$\begin{aligned} \sup_{y \geq 2^{m-3/2}} (4y^2 + 2^{-m})P(Z \leq \lambda - 2^{m/2}y) &= (2^{2m-1} + 2^{-m})P(Z \geq 2^{3(m-1)/2} - \lambda) \\ &\leq \frac{2^{2m-1} + 2^{-m}}{2^{3(m-1)/2} - \lambda}\phi(2^{3(m-1)/2} - \lambda), \end{aligned}$$

where $\phi(\cdot)$ is the density function of the standard normal distribution. It is also easy to check that $(4y^2 + 2^{-2})P(Z \leq \lambda - 2y)$ is also decreasing in y and so

$$\sup_{y \geq \sqrt{2}} (4y^2 + 2^{-2})P(Z \leq \lambda - 2y) = \frac{33}{4}P(Z \geq \sqrt{2}) < 0.649.$$

For $m = 1$, (74) can be verified through direct numerical calculations. ∎

Finally, we state and prove the following result which was used in the proof of Lemma 6. This lemma is useful in obtaining a lower bound for the local modulus of continuity. It is helpful to first plot the functions involved.



Lemma 11 *Fix $a > 0$. Let $u > 0$ and $0 \le v < 1$. Define $f_u(t) = -a + ut$ and $g_v(t) = \frac{a}{1-v}(t - v) \cdot I(t \ge v)$. Then*

$$\inf_{u \ge 0, 0 \le v < 1} \int_0^1 (f_u(t) - g_v(t))^2 dt \ge 0.1572a^2. \tag{77}$$

Proof of Lemma 11: Set $S = \int_0^1 (f_u(t) - g_v(t))^2 dt$. Then

$$\begin{aligned} S &= \int_0^v (-a+ut)^2 dt + \int_v^1 (-a + ut + \frac{av}{1-v} - \frac{at}{1-v})^2 dt \\ &= \int_0^v (a^2 - 2aut + u^2t^2) dt \\ &\quad + \frac{1}{(1-v)^2} \int_v^1 \left((u - uv - a)^2 t^2 + 2a(2v-1)(u-uv-a)t + a^2(2v-1)^2\right) dt \\ &= a^2 v - auv^2 + \frac{1}{3}u^2v^3 + \frac{1}{1-v} \cdot \Delta \end{aligned}$$

where

$$\Delta = \frac{1}{3}(u - uv - a)^2(1 + v + v^2) + a(2v-1)(u - uv - a)(1 + v) + a^2(2v-1)^2.$$

We shall first simplify Δ. Setting $w = 1 - v$. Tedious but straightforward algebra shows that

$$\Delta = w\left\{\frac{1}{3}(u^2w - 2au)(w^2 - 3w + 3) + \frac{7}{3}a^2 w + au(2w^2 - 5w + 2)\right\}.$$

Combining this with other terms yields

$$\begin{aligned} S &= a^2(1-w) - au(1-w)^2 + \frac{1}{3}u^2(1-w)^3 \\ &\quad + \frac{1}{3}(u^2w^3 - (2au + 3u^2)w^2 + (6au + 3u^2)w - 6au) \\ &\quad + 2auw^2 + (\frac{7}{3}a^2 - 5au)w + 2au \\ &= \frac{1}{3}auw^2 + (\frac{4}{3}a^2 - au)w + (a^2 - au + \frac{1}{3}u^2). \end{aligned}$$

Note that a is fixed and so S is a function of u and w. We wish to minimize $S = S(u, w)$ with respect to u and w. Setting the partial derivatives to 0, we have

$$\left\{ \begin{array}{rcl} \frac{2}{3}auw + \frac{4}{3}a^2 - au &=& 0 \\ \frac{1}{3}aw^2 - aw - a + \frac{2}{3}u &=& 0 \end{array} \right. .$$

This yields

$$\left\{ \begin{array}{c} u = (-\frac{1}{2}w^2 + \frac{3}{2}w + \frac{3}{2})a \\ 2w^3 - 9w^2 + 3w + 1 = 0 \end{array} \right. .$$

The cubic equation has a unique root between 0 and 1, $w = 0.5986$ and the corresponding value of u is $u = 2.2187a$. The minimum value of S is $S = S(2.2187, 0.5986) = 0.1572a^2$. ∎