

Stability Properties of Perona-Malik Scheme

Selim Esedoglu

Institute for Mathematics and its Applications

December 2001

Abstract

The Perona-Malik scheme is a numerical technique for de-noising digital images without blurring object boundaries (edges). In general, solutions generated by this scheme do not satisfy a comparison principle. We identify conditions under which two solutions initially ordered remain ordered, and state (restricted) comparison principles. These allow us to study stability properties of the scheme. We also explore what these results say in the limit as the discretization size goes to 0.

1. Introduction

In this paper we state a *restricted* comparison principle for the semidiscrete Perona-Malik scheme. This scheme is described by the system of o.d.e.'s

$$\dot{u}_{i,j}(t) = D_1^- (R_\kappa (D_1^+ u_{i,j})) + D_2^- (R_\kappa (D_2^+ u_{i,j})) \quad (1)$$

where D_m^+ and D_m^- are the standard forward and backward difference operators in the m -th coordinate direction, and the function R_κ satisfies important properties which we explain in Section 2. We note that a *restriction* of the kind we consider is necessary, because the standard comparison result does not hold for this scheme. We illustrate this easy to see fact by a simple numerical example (*Figure 2*).

Subsequently, we apply the comparison principle to explore stability properties of the scheme at fixed discretization step size $h > 0$ and also in the limit $h \rightarrow 0^+$. We also expose some effects the precise shape of the function R_κ has on the behavior of the scheme.

Scheme (1) was proposed by Perona and Malik as a numerical technique for de-noising digital images without blurring object boundaries [6, 7]. In particular, it is an alternative to Gaussian smoothing, which is equivalent to solving the heat equation with the original image as initial data. The motivation was to replace the standard heat equation by the following non-linear diffusion equation which is designed to suppress diffusion at regions of large gradient, and for which scheme (1) is a natural finite differences discretization:

$$u_t = (R_\kappa(u_x))_x + (R_\kappa(u_y))_y \quad (2)$$

The essence of the technique is contained in the choice of the function $R_\kappa(\xi)$; the success of the scheme in preserving sharp edges (until they disappear) is due to this choice. But for the functions R_κ that Perona and Malik advocate in their papers, p.d.e. (2) becomes backwards parabolic in regions where the gradient of the solution is larger than some threshold that depends on the parameter κ . As such, there is no well-posedness theory for this pde. That makes it interesting to investigate continuum limits (limit as $h \rightarrow 0$) for scheme (1). And understanding how the behavior of a de-noising technique such as (1) depends on the discretization step size is important since it is very common to have the same image at various resolutions.

In practice, scheme (1) exhibits much better stability properties than one would expect from backwards diffusion equations, which are notoriously ill-posed. Moreover, it is extremely effective at its intended purpose. It has therefore become an intriguing issue to explain the better than expected stability properties of Perona and Malik's technique. Some aspects of this surprisingly tame behavior have been explained by previous authors; we believe with this paper we further our understanding of this problem.

Another interesting issue is what effects the precise shape of the non-linear function R_κ has on the scheme. Perona and Malik, and subsequently many other authors, reported numerical experiments with a variety of choices (each of which conforms to the fundamental properties we listed in Section 2), and on occasion mentioned differences in observed behavior [8]. Indeed, based on numerical experiments, even with functions R_κ that have identical parabolicity thresholds the behavior can still be quite different. The results presented in this paper allow us to reveal and quantify some differences.

2. Perona-Malik Scheme

As we remarked above, it is the choice of $R_\kappa(\xi)$ that distinguishes Perona and Malik's technique from previous ones. In [6], they report numerical experiments using scheme (1) with

$$R_\kappa(\xi) = \frac{\xi}{1 + \xi^2/\kappa} \text{ and } R_\kappa(\xi) = \xi \exp\left(\frac{-\xi^2}{\kappa}\right)$$

Other choices used in practice include

$$R_\kappa(\xi) = \xi \left(1 + \frac{\xi^2}{\kappa}\right)^{(\beta-1)} \text{ where } \beta \in \left(0, \frac{1}{2}\right)$$

These choices share the following essential characteristics:

1. $\xi R_\kappa(\xi) \geq 0$ for all ξ .

2. The parameter κ defines a positive critical value $z(\kappa)$ such that

$$R'_\kappa(\xi) \begin{cases} < 0 \text{ for } |\xi| > z(\kappa), \text{ and} \\ \geq 0 \text{ otherwise.} \end{cases} \quad (3)$$

3. $R_\kappa(\xi) \rightarrow 0$ as $|\xi| \rightarrow \infty$.

Figure 1 illustrates $R_\kappa(\xi)$ for such a choice.

In (2), R'_κ appears as the diffusion coefficient. Therefore, as we indicated in the previous section, the parameter κ constitutes a threshold value: when the gradient of grayscale intensity, $D_m^+ u_{i,j}$, is large compared to κ , equation (2) violates parabolicity.

Encouraged by favorable numerical results, some previous mathematical work on the Perona-Malik technique deals with understanding whether equation (2) can be given a well-posedness theory, so that the p.d.e. (2) can be properly understood as the continuum limit of scheme (1). The paper [5] by Kichenassamy, and paper [4] by Kawohl and Kutev pursue this direction. The present work is drastically different from them in spirit: we do not try to understand equation (2) at all; instead, we deal directly with scheme (1) and study its properties. This is also the approach pursued in [2], where a continuum limit for (1) is established that differs from (2).

Nevertheless, we make use of ideas from the work of these previous researchers; indeed, the motivation for this paper came from [4]. There, Kawohl and Kutev establish a restricted comparison principle for *continuum* solutions of (2). In this paper, we strive to find conditions under which two *discrete* solutions generated by the scheme (1) can be compared. In the end, however, we did not obtain direct discrete analogues of Kawohl and Kutev's results; the conditions and results in this paper are entirely different.

3. Comparison Principle

We begin by introducing some notation. First, from now on the subscript κ in R_κ will be suppressed, but it will be understood that R comes with a parameter κ . Notice that R is a two-to-one function; we will denote by R_1 and R_2 the restriction of R to the domains $[-z(\kappa), z(\kappa)]$ and $(-z(\kappa), z(\kappa))^c$ respectively:

$$R_1(\xi) := R(\xi)|_{[-z(\kappa), z(\kappa)]}, \text{ and } R_2(\xi) := R(\xi)|_{(-z(\kappa), z(\kappa))^c}$$

Then, R_1 and R_2 are one-to-one functions, whose inverses will be denoted R_1^{-1} and R_2^{-1} respectively.

We will speak of the ‘‘jump set’’ $S(\phi)$ of a function $\{\phi_{i,j}\}$ defined on the grid; with that we mean the collection of indices defined by

$$S(\phi) := \{(i, j) : \max(|D_1^+ \phi_{i,j}|, |D_2^+ \phi_{i,j}|) \geq z(\kappa)\}$$

Also, we adopt the terminology in [4] to say that ϕ is *supersonic* on $S(\phi)$, and *subsonic* elsewhere.

Proposition 1 *Let $\{u_{i,j}(t)\}$ and $\{v_{i,j}(t)\}$ be solutions generated by the Perona-Malik scheme (1), subject to Neumann boundary conditions. Assume that:*

1. $|D_m^+ v_{i,j}(t)| < z(\kappa)$ for all $(i, j), t \in [0, T]$, and $m = 1, 2$,
2. $|D_m^+ u_{i,j}(t)| \leq R_2^{-1} (R(|D_m^+ v_{i,j}(t)|))$ for all $(i, j), t \in [0, T]$, and $m = 1, 2$.

Then if $\{u_{i,j}(t)\}$ and $\{v_{i,j}(t)\}$ are strictly ordered at $t = 0$, they remain ordered for all $t \in [0, T]$; i.e.

1. *if $u_{i,j}(0) > v_{i,j}(0)$ then $u_{i,j}(t) \geq v_{i,j}(t)$ for all $t \in [0, T]$,*
2. *if $u_{i,j}(0) < v_{i,j}(0)$ then $u_{i,j}(t) \leq v_{i,j}(t)$ for all $t \in [0, T]$.*

Proof: We only treat the first case $u_{i,j}(0) > v_{i,j}(0)$, the second case being completely analogous. Suppose the conclusion is false. Then there exists $t_0 \in (0, T]$ such that

$$\begin{aligned} u_{i,j}(t) &> v_{i,j}(t) \text{ for all } (i, j) \text{ and } t \in [0, t_0], \text{ and} \\ u_{k,l}(t_0) &= v_{k,l}(t_0) \text{ for some } (k, l). \end{aligned}$$

Consequently, $\dot{v}_{k,l}(t_0) - \dot{u}_{k,l}(t_0) \geq 0$, and hence

$$\begin{aligned} D_1^- R(D_1^+ v_{k,l}(t_0)) - D_1^- R(D_1^+ u_{k,l}(t_0)) \\ + D_2^- R(D_2^+ v_{k,l}(t_0)) - D_2^- R(D_2^+ u_{k,l}(t_0)) \geq 0 \end{aligned}$$

Therefore,

$$\text{Either } D_1^- R(D_1^+ v_{k,l}(t_0)) - D_1^- R(D_1^+ u_{k,l}(t_0)) \geq 0, \quad (4a)$$

$$\text{or } D_2^- R(D_2^+ v_{k,l}(t_0)) - D_2^- R(D_2^+ u_{k,l}(t_0)) \geq 0. \quad (4b)$$

Without loss of generality, assume that (4a) is true. That means:

$$\begin{aligned} R(D_1^+ v_{k,l}(t_0)) - R(D_1^+ u_{k,l}(t_0)) \\ + R(D_1^+ u_{k-1,l}(t_0)) - R(D_1^+ v_{k-1,l}(t_0)) \geq 0 \end{aligned} \quad (5)$$

But now, by hypothesis 1 and 2,

$$\left(R(D_1^+ v_{k,l}(t_0)) - R(D_1^+ u_{k,l}(t_0)) \right) \left(D_1^+ v_{k,l}(t_0) - D_1^+ u_{k,l}(t_0) \right) \geq 0 \quad (6a)$$

and,

$$\begin{aligned} & \left(R(D_1^+ u_{k-1,l}(t_0)) - R(D_1^+ v_{k-1,l}(t_0)) \right) \\ & \quad \times \left(D_1^+ u_{k-1,l}(t_0) - D_1^+ v_{k-1,l}(t_0) \right) \geq 0 \quad (6b) \end{aligned}$$

By definition of t_0 , we also have:

$$\begin{aligned} D_1^+ v_{k,l}(t_0) - D_1^+ u_{k,l}(t_0) &= D_1^+ (v_{k,l} - u_{k,l})(t_0) \leq 0, \text{ and} \\ D_1^+ u_{k-1,l}(t_0) - D_1^+ v_{k-1,l}(t_0) &= D_1^+ (u_{k-1,l} - v_{k-1,l})(t_0) \leq 0 \end{aligned} \quad (7)$$

So we get, in particular:

$$\begin{aligned} & \left(R(D_1^+ v_{k,l}(t_0)) - R(D_1^+ u_{k,l}(t_0)) \right) \\ & \quad \times \left(R(D_1^+ u_{k-1,l}(t_0)) - R(D_1^+ v_{k-1,l}(t_0)) \right) \geq 0 \end{aligned}$$

By (5),

$$\begin{aligned} R(D_1^+ v_{k,l}(t_0)) - R(D_1^+ u_{k,l}(t_0)) &\geq 0 \\ R(D_1^+ u_{k-1,l}(t_0)) - R(D_1^+ v_{k-1,l}(t_0)) &\geq 0 \end{aligned}$$

By (6a) and (6b) that means:

$$\begin{aligned} D_1^+ v_{k,l}(t_0) - D_1^+ u_{k,l}(t_0) &\geq 0 \\ D_1^+ u_{k-1,l}(t_0) - D_1^+ v_{k-1,l}(t_0) &\geq 0 \end{aligned} \quad (8)$$

Finally, (7) and (8) imply that

$$D_1^+ v_{j,l}(t_0) = D_1^+ u_{j,l}(t_0) \text{ for } j = k-1, k$$

and thus:

$$v_{j,l}(t_0) = u_{j,l}(t_0) \text{ for } j = k-1, k, k+1.$$

As a result, equality holds in (4a). Therefore, (4b) is also true. The same line of reasoning we used for (4a) now gives:

$$v_{k,j}(t_0) = u_{k,j}(t_0) \text{ for } j = l-1, l, l+1.$$

Repetition of this argument (with k replaced by $k \pm 1$ and l replaced by $l \pm 1$, and so on) gives:

$$v_{i,j}(t_0) = u_{i,j}(t_0) \text{ for all } (i, j).$$

But then uniqueness of the solution to system (1) implies that

$$v_{i,j}(t) = u_{i,j}(t) \text{ for all } (i, j) \text{ and } t \geq t_0$$

which proves the proposition. \square

Remark: Proposition 1 has been stated and proved only on a two-dimensional grid for simplicity; the statement is true, and the proof works with small modifications, for any space dimension.

Proposition 1 allows for comparison when one of the solutions is “smooth” (i.e. subsonic). In the case of one space dimension, we shall say a bit more: the next proposition allows for the comparison of more general one-dimensional signals. Its proof is a slight variation on that of Proposition 1. Its hypothesis will be justified in the next section, especially through Proposition 3. And eventually, it will find an application in Section 4.3, where we will consider the behavior of scheme (1) as $h \rightarrow 0^+$.

Proposition 2 *Let $\{u_j(t)\}$ and $\{v_j(t)\}$ be one-dimensional solutions generated by the Perona-Malik scheme (1), subject to Neumann boundary conditions. Assume that:*

1. $S(v(t)) = S(v(0)) := \{p_1, \dots, p_n\} \subseteq S(u(t))$ for all $t \in [0, T]$,
2. $|D^+u_j(t)| \leq R_2^{-1}(R(|D^+v_j(t)|))$ for all $j \notin S(v(0))$ and $t \in [0, T]$,
3. $\text{sign}(u_i(0) - v_i(0)) = \text{sign}(u_j(0) - v_j(0))(-1)^{k-k'} \neq 0$ for $i \in \{p_k + 1, \dots, p_{k+1}\}$ and $j \in \{p_{k'} + 1, \dots, p_{k'+1}\}$.

Then, for all $t \in [0, T]$ we have:

$$(u_j(t) - v_j(t))(u_j(0) - v_j(0)) \geq 0$$

Proof: The conclusion is satisfied for some positive time by continuity; suppose that it fails for the first time at $t = t_0 < T$ and at index k . Without loss of generality, let us assume that $u_k(0) > v_k(0)$. Define $\alpha : \mathbf{Z} \rightarrow \{0, 1\}$ as follows:

$$\alpha(j) := \begin{cases} 1 & \text{if } j \in S(v(0)) \\ 0 & \text{otherwise.} \end{cases}$$

By definitions of t_0 and k , and by hypothesis 3, we have:

$$\begin{aligned} (D^+u_k(t_0) - D^+v_k(t_0)) (-1)^{\alpha(k)} &\geq 0 \\ (D^+u_{k-1}(t_0) - D^+v_{k-1}(t_0)) (-1)^{\alpha(k-1)} &\leq 0 \end{aligned} \tag{9}$$

Hypothesis 2 implies, as in the proof of Proposition 1, that

$$\left(R(D^+u_j(t)) - R(D^+v_j(t))\right)\left(D^+u_j(t) - D^+v_j(t)\right) \geq 0 \text{ if } j \notin S(v(0)) \quad (10)$$

On the other hand, if $j \in S(v(0))$, then $|D^+u_j|, |D^+v_j| \geq z(\kappa)$; and since R is decreasing on $(-z(\kappa), z(\kappa))^c$ we get:

$$\left(R(D^+u_j(t)) - R(D^+v_j(t))\right)\left(D^+u_j(t) - D^+v_j(t)\right) \leq 0 \text{ if } j \in S(v(0)) \quad (11)$$

We can summarize (10) and (11) as

$$\left(R(D^+u_j(t)) - R(D^+v_j(t))\right)\left(D^+u_j(t) - D^+v_j(t)\right)(-1)^{\alpha(j)} \geq 0 \quad (12)$$

Furthermore, the definition of t_0 implies that

$$\begin{aligned} \dot{u}_k(t_0) - \dot{v}_k(t_0) &= R(D^+u_k(t_0)) - R(D^+v_k(t_0)) \\ &\quad - R(D^+u_{k-1}(t_0)) + R(D^+v_{k-1}(t_0)) \\ &\leq 0 \end{aligned} \quad (13)$$

But now, (9) and (12) mean that

$$\begin{aligned} &\left(R(D^+u_k(t_0)) - R(D^+v_k(t_0))\right) \\ &\quad \times \left(R(D^+v_{k-1}(t_0)) - R(D^+u_{k-1}(t_0))\right) \geq 0 \end{aligned} \quad (14)$$

Combined with (13), (14) implies:

$$\begin{aligned} R(D^+u_k(t_0)) - R(D^+v_k(t_0)) &\leq 0 \\ R(D^+v_{k-1}(t_0)) - R(D^+u_{k-1}(t_0)) &\leq 0 \end{aligned} \quad (15)$$

In light of (12), these inequalities lead to the conclusion:

$$\begin{aligned} (D^+u_k(t_0) - D^+v_k(t_0))(-1)^{\alpha(k)+1} &\geq 0 \\ (D^+u_{k-1}(t_0) - D^+v_{k-1}(t_0))(-1)^{\alpha(k-1)} &\geq 0 \end{aligned} \quad (16)$$

But then, (16) and (9) give:

$$D^+(u_j(t_0) - v_j(t_0)) = 0 \text{ for } j = k-1, k$$

so that

$$u_j(t_0) = v_j(t_0) \text{ for } j = k-1, k, k+1$$

That, much like in the proof of Proposition 1, leads to the conclusion of the Proposition. \square

4. Applications

The comparison principles stated and proved in the previous section are simply tools; indeed, their hypothesis require knowledge of the solutions involved for all time. Here, in Section 4.1, we use them to state some down to earth results, such as the stability property that is the content of Theorem 1. Then, in Section 4.2, we give concrete examples of how those results can be applied in practice. Section 4.3 is devoted to exploring what these results say in the limit as $h \rightarrow 0^+$.

4a. Stability Results

We begin by recording a few simple but important properties of scheme (1) that will help us apply Propositions 1 and 2.

Lemma 1 *Let $\{u_{i,j}(t)\}$ be the solution generated by scheme (1) from subsonic initial data. Then*

$$\sup_{i,j,t} |D_m^+ u_{i,j}(t)| \leq \sup_{i,j} |D_m^+ u_{i,j}(0)| \text{ for each } m = 1, 2.$$

Proof: It is easy to see that in the subsonic regime, scheme (1) satisfies a maximum principle for difference quotients. This in turn prevents the solution from entering the supersonic regime, if the initial data is subsonic. The conclusion of the lemma follows. \square

In what follows, we will often specialize to the one-dimensional version of the Perona–Malik scheme (1), which then reduces to:

$$\dot{u}_j(t) = D^- (R_\kappa (D^+ u_j)) \tag{17}$$

Next, we recall an important property of scheme (17): supersonic regions shrink in time.

Proposition 3 *Let $\{u_j(t)\}$ be a solution generated by scheme (17). Then $S(u(t_2)) \subseteq S(u(t_1))$ whenever $0 \leq t_1 \leq t_2$.*

Proof: See [3] where it first appeared, or [2].

Remark: The conclusion of Proposition 3 is false in two-dimensions: supersonic regions can grow, as shown by the numerical experiment in *Figure 3*.

Our first application deals with subsonic data corrupted by low amplitude noise. We estimate the difference between the evolutions of corrupted and uncorrupted data in terms of the amplitude of the noise. The hypothesis of Proposition 1 involve

all $t \geq 0$. We will use in our proof sub and super solutions based only on the initial data. As we shall explain, they will satisfy the hypothesis automatically for all time.

Theorem 1 *Let $\{\phi_{i,j}\}$ be subsonic initial data, i.e.*

$$M := \max_{i,j,m} |D_m^+ \phi_{i,j}| < z(\kappa)$$

Let $\{u_{i,j}(t)\}$ be the solution generated by scheme (1) from $\{\phi_{i,j}\}$, and let $\{u_{i,j}^n(t)\}$ be the one generated from $\{(\phi + n)_{i,j}\}$. If

$$\max_{i,j} |n_{i,j}| < \frac{h}{2} (R_2^{-1}(R(M)) - M),$$

then

$$\max_{i,j} |u_{i,j}^n(t) - u_{i,j}(t)| \leq \max_{i,j} |n_{i,j}| \text{ for all } t \geq 0.$$

Proof: Fix a $\delta > 0$ such that

$$\max_{i,j} |n_{i,j}| < \delta < \frac{h}{2} (R_2^{-1}(R(M)) - M).$$

The sub and super solutions, which we shall denote $v_{i,j}^-(t)$ and $v_{i,j}^+(t)$ respectively, will simply be:

$$v_{i,j}^\pm(t) := u_{i,j}(t) \pm \delta \tag{18}$$

Then $v_{i,j}^\pm(t)$ are clearly solutions of (1). Furthermore, $u_{i,j}^n(0) \in (v_{i,j}^-(0), v_{i,j}^+(0))$. Since the initial condition $\phi_{i,j}$ is subsonic, by Lemma 1, $u_{i,j}(t)$ and therefore also $v_{i,j}^\pm(t)$ are subsonic for all time. Thus, *Hypothesis 1* in Proposition 1 is satisfied. Moreover, again by virtue of Lemma 1, we have:

$$\max_{i,j,m,t} |D_m^+ u_{i,j}(t)| = \max_{i,j,m,t} |D_m^+ v_{i,j}^\pm(t)| = M \tag{19}$$

Also, inequality in *Hypothesis 2* of Proposition 1 is *strictly* satisfied at $t = 0$ since

$$\begin{aligned} |D_m^+ u_{i,j}^n(0)| &\leq |D_m^+ u_{i,j}(0)| + \frac{2}{h} \max_{i,j} |n_{i,j}| \\ &< R_2^{-1}(R(M)) \leq R_2^{-1}(R(|D_m^+ u_{i,j}(0)|)) \end{aligned}$$

by our assumption on the amplitude of the noise $n_{i,j}$. We will now show that in fact *Hypothesis 2* is strictly satisfied for all time. Suppose not; then there exists $t_0 > 0$ such that

$$|D_m^+ u_{i,j}^n(t)| < R_2^{-1} (R (|D_m^+ u_{i,j}(t)|))$$

for all (i, j) and $m \in \{1, 2\}$ and $t \in [0, t_0)$, and

$$|D_m^+ u_{k,l}^n(t_0)| = R_2^{-1} (R (|D_m^+ u_{k,l}(t_0)|))$$

for some (k, l) and some $m \in \{1, 2\}$. By (19), that means

$$|D_m^+ u_{k,l}^n(t_0)| \geq R_2^{-1} (R(M)) \quad (20)$$

We also have

$$\begin{aligned} |D_m^+ u_{k,l}^n(t_0)| &\leq |D_m^+(u^n - u)_{k,l}(t_0)| + |D_m^+ u_{k,l}(t_0)| \\ &\leq |D_m^+(u^n - u)_{k,l}(t_0)| + M \end{aligned} \quad (21)$$

again by (19). Combining (20) and (21) we get

$$|D_m^+(u^n - u)_{k,l}(t_0)| \geq R_2^{-1} (R(M)) - M > \frac{2\delta}{h}$$

That means we have

$$|u_{i,j}^n(t_0) - u_{i,j}(t_0)| > \delta \text{ for some } (i, j) \in \{k, k+1\} \times \{l, l+1\} \quad (22)$$

On the other hand, since both hypothesis of Proposition 1 are satisfied on $t \in [0, t_0)$, we have

$$v_{i,j}^-(t) \leq u_{i,j}^n(t) \leq v_{i,j}^+(t) \text{ for all } t \in [0, t_0)$$

and, by continuity, also at $t = t_0$. Combined with (18) this means

$$|u_{i,j}^n(t_0) - u_{i,j}(t_0)| \leq \delta \text{ for all } (i, j)$$

which contradicts (22). \square

An immediate consequence of Theorem 1 is the following elementary corollary, which is, unlike the theorem, one dimensional. It tells us that a smooth one dimensional signal corrupted by low amplitude noise is rapidly de-noised, and provides an upper bound on the de-noising time:

Corollary 1 *Let ϕ_j, n_j, u_j, u_j^n , and M be as in Theorem 1, and assume that n_j satisfies the hypothesis of that theorem. If we set*

$$T := \inf\{t_0 \geq 0 : S(u^n) \text{ is empty for all } t \geq t_0\}$$

then we have the estimate:

$$T \leq \frac{2 \max_j |\phi_j + n_j|}{R\left(\frac{2}{h} \max_j |n_j| + M\right)}$$

Proof: The interesting case is when $S(u^n(0))$ is non-empty; under that assumption, for any $\delta > \max_j |n_j|$ we have $2\delta/h + M > z(\kappa)$. Fix an $\varepsilon > 0$ small enough so that:

$$R(z(\kappa) - \varepsilon) > R\left(\frac{2\delta}{h} + M\right)$$

For $k \in S(u^n(0))$, let

$$T_k^\varepsilon := \inf\{t \geq 0 : |D^+ u_k^n(t)| = z(\kappa) - \varepsilon\}$$

In view of Proposition 3, we have

$$T < \max_k T_k^\varepsilon$$

So fix a $k \in S(u^n(0))$; without loss of generality, we may assume that $D^+ u_k^n(0) > 0$. By Theorem 1,

$$D^+ u_k^n(t) < \frac{2\delta}{h} + M \text{ for all } t \geq 0.$$

Furthermore, since R is a decreasing function on $[z(\kappa), \infty)$, for $t \in [0, T_k^\varepsilon]$ we have:

$$z(\kappa) - \varepsilon \leq D^+ u_k^n(t) < \frac{2\delta}{h} + M \implies R(D^+ u_k^n(t)) > R\left(\frac{2\delta}{h} + M\right) \quad (23)$$

That gives:

$$\frac{d}{dt} \sum_{j=1}^k h \left(\max_j |\phi_j + n_j| - u_j^n \right) = -R(D^+ u_k^n) \leq -R\left(\frac{2\delta}{h} + M\right)$$

where the inequality follows via (23) for as long as $t \in [0, T_k^\varepsilon]$. Also, since $|u_k^n(t)| \leq \max_j |\phi_j + n_j|$ for all $t \geq 0$, we have:

$$2(\max_j |\phi_j + n_j|) - T_k^\varepsilon R\left(\frac{2\delta}{h} + m\right) \geq 0$$

Letting $\delta \rightarrow \max_j |n_j|$ from above leads to the desired inequality. \square

4b. Examples

In this section, we apply the results of the previous subsection in some practical situations.

Example 1: Take the function in *Figure 4*. It has maximum slope $M = 4$. The Perona-Malik scheme (17) is applied using the choice:

$$R(\xi) = \frac{\xi}{1 + \frac{\xi^2}{100}} \quad (24)$$

for the nonlinear function appearing in the scheme. According to this choice, the threshold value of slope is $z(\kappa) = 10$. We calculate the maximum noise amplitude allowed by Theorem 1 to be 0.0525. The corrupted signal in the example of *Figure 4*, which is not subsonic, was obtained by adding noise of amplitude 0.05 to the original signal.

The evolution shown in *Figure 5* is obtained by adding a specific perturbation of amplitude 0.06 to the original signal. We see how the comparison principle gets violated.

Example 2: We now compare the effects of the precise shape of the function $R(\xi)$ on the behavior of the scheme, by using the same original image as in our first example, but the different nonlinear function

$$R(\xi) = \xi \exp\left(-\frac{\xi^2}{200}\right) \quad (25)$$

which has the same threshold value of the slope as for (24) of the first example, namely $z(\kappa) = 10$. The maximum amplitude of noise allowed by Theorem 1 this time (for $R(\xi)$ given by (25)) turns out to be between 0.034 and 0.03425 – significantly smaller than that for (24).

Theorem 1 thus suggests a way to quantify the difference in stability properties of scheme (1) with respect to the two choices of $R_\kappa(\xi)$. It is easy to see that this difference is important; we illustrate it with a numerical example: *Figure 6* shows the evolution of the original image perturbed by a (contrived) noise of amplitude 0.05 under scheme (17) using the two choices for $R(\xi)$ given in (24) and (25). Note that 0.05 is above the allowed limit for (25), and below it for (24).

4c. Limit as $h \rightarrow 0^+$

In [2], the continuum limit of scheme (17) is investigated with the function $R_\kappa(\xi)$ given by

$$R_\kappa(\xi) = \xi \left(1 + \frac{\xi^2}{\kappa}\right)^{(\beta-1)} \quad \text{where } \beta \in \left[0, \frac{1}{2}\right). \quad (26)$$

When $\beta \in (0, \frac{1}{2})$ the evolution that (17) generates is the gradient descent for the discrete energy

$$\mathbf{E}_u^h(t) := \sum_j h \Phi_{\kappa, \beta} ((D^+ u_j(t))^2) \quad (27)$$

where

$$\Phi_{\kappa, \beta}(\xi) := \frac{\kappa}{\beta} \left(\left(1 + \frac{\xi}{\kappa} \right)^\beta - 1 \right) \quad (28)$$

The continuum limit studied in [2] is obtained by scaling the parameter κ (and hence the threshold point $z(\kappa)$) with respect to the discretization step size h as follows:

$$\kappa(h) = h^{(2\beta-1)/(1-\beta)} \implies z(h) := \frac{1}{\sqrt{1-2\beta}} h^{(2\beta-1)/(2-2\beta)} \quad (29)$$

Such scalings were studied previously in the stationary setting by Chambolle in [1] to obtain interesting continuum limits for discrete energies similar to (27); the approach taken in [2] follows Chambolle's lead in adjusting the threshold z with respect to the grid size h , but concerns the time dependent problem. The resulting evolution is defined for piecewise smooth one-dimensional signals; it takes place on a domain that changes at discrete times, and is described as follows:

For a given piecewise smooth initial data $\phi : [0, 1] \rightarrow [0, 1]$ with jump discontinuities at $p_1, \dots, p_N \in (0, 1)$, we solve the linear heat equation $u_t = u_{xx}$ on the domain $(0, 1) - \{p_1, \dots, p_N\}$, subject to homogeneous Neumann boundary conditions at $x = 0$ and $x = 1$, and to the nonlinear boundary conditions

$$u_x(p_j^\pm, t) = \left(u(p_j^+, t) - u(p_j^-, t) \right) \left| u(p_j^+, t) - u(p_j^-, t) \right|^{2\beta-2} \quad (30)$$

at the discontinuity points. The condition (30) becomes singular whenever one of the discontinuities, say the one at p_j , heals (i.e. when $u(p_j^+, t) = u(p_j^-, t)$); at such special times, we merge the two intervals (p_{j-1}, p_j) and (p_j, p_{j+1}) into one longer interval (p_{j-1}, p_{j+1}) , and continue the evolution according to the heat equation on the new (and smaller) collection of intervals.

The claim is that the numerical solutions generated by scheme (17) converge, as $h \rightarrow 0$, to the continuum evolution described above, provided that the threshold $z(\kappa)$ is scaled with respect to h according to formula (29), and that the approximate (discrete) initial data ϕ^h converge to the continuum data ϕ in some suitable sense. The proof in [2] for this statement involves a number of technical hypothesis. One of the most restrictive among them requires the ‘‘jump sets’’ of ϕ^h and ϕ to be *compatible*: it is assumed that $S(\phi^h)$ and $\{p_1, \dots, p_N\}$ are in one-to-one correspondence.

A discussion of the most general conditions under which convergence takes place would be very technical and out of place. But, as an application of Proposition 2, we will show that under suitable hypothesis, the jump sets $S(\phi^h)$ of the approximate initial data become compatible with $\{p_1, \dots, p_N\}$ after an arbitrarily small initial interval of time. To that end, let $\{h_n\}_{n=1}^\infty$ be a sequence of positive numbers such that $h_n \rightarrow 0$ as $n \rightarrow \infty$, and let $x_j^{h_n}$ denote the grid points for the uniform discretization size h_n . Assume that a sequence ϕ^{h_n} of discrete initial data satisfies

$$\lim_{n \rightarrow \infty} \max_j \left| \phi_j^{h_n} - \phi(x_j^{h_n}) \right| = 0$$

Then we have the following result:

Theorem 2 *Let $\{u^{h_n}(t)\}_n$ be the discrete solutions generated from $\{\phi^{h_n}\}_n$ by scheme (17), where R is given by (26) and κ is scaled as in (29). Then, given any $\varepsilon > 0$, there exists $K \in \mathbf{N}$ such that for any $n > K$ the following property holds at some $t \in [0, \varepsilon)$:*

$$|D^+ u_j^{h_n}(t)| \geq z(\kappa) \text{ only if } \{p_1, \dots, p_N\} \cap [x_j^{h_n}, x_{j+1}^{h_n}] \text{ is non-empty.}$$

Proof: The jump sets of initial data ϕ^{h_n} can be much larger than that of ϕ ; the idea is to construct comparison functions $\psi^{n,\pm}$ whose jump sets precisely match $\{p_1, \dots, p_N\}$, and then compare using Proposition 2. We first define:

$$M := \max_i \max_{x \in (p_i, p_{i+1})} |\phi'(x)|, \quad m := \min_i |\phi(p_i^+) - \phi(p_i^-)|$$

$$\tilde{S}_n := \left\{ j \in \mathbf{N} : \left| D^+ \phi_j^{h_n} \right| \geq \frac{m}{2h_n} \right\}, \quad \alpha(j) := \#\{i \in \tilde{S}_n : i < j\}$$

Convergence of the approximate initial data ϕ^{h_n} to ϕ uniformly on the grid implies that for large enough n we have

$$\begin{aligned} j \in \tilde{S}_n &\implies p_i \in [x_j, x_{j+1}] \text{ for some } i, \text{ and} \\ p_i \in [x_j, x_{j+1}] &\implies \{j, j+1\} \cap \tilde{S}_n \neq \emptyset \end{aligned} \tag{31}$$

For $\delta \in (0, \frac{m}{4})$, we construct the pair of comparison functions $\psi^{n,\pm}(t)$ via scheme (17) from the following initial data:

$$\psi_j^{n,\pm}(0) = \phi(x_j^{h_n}) \pm \delta(-1)^{\alpha(j)}$$

Then, $\psi^{n,\pm}(0)$ satisfy the following properties:

$$S(\psi^{n,\pm}(0)) = \tilde{S}(\phi^{h_n}) \tag{32a}$$

$$\#\mathcal{S}(\psi^{n,\pm}(0)) = N, \text{ and} \tag{32b}$$

$$\sup_n \mathbf{E}_{\psi^{n,\pm}}^{h_n}(0) < \infty \tag{32c}$$

Furthermore,

$$\limsup_{n \rightarrow \infty} \max_{j \notin S(\psi^{n,\pm})} |D^+ \psi^{n,\pm}| = M < \infty \quad (32d)$$

We recall a few points from [2]: First, by virtue of property (32c), the evolutions $\{\psi^{n,\pm}(t)\}_n$ are Holder continuous in time with values in L^∞ of space, *uniformly* in n . Moreover, the difference quotients of $\psi^{n,\pm}(t)$ satisfy the maximum principle on the complements of their jump sets, while the jump sets remain constant. In light of these comments and of Proposition 3, we can determine a $T > 0$ so that for all $t \in [0, T]$ the following hold:

$$S(\psi^{n,\pm}(t)) = S(\psi^{n,\pm}(0)) \quad (33a)$$

$$\max_{j \notin S(\psi^{n,\pm})} |D^+ \psi^{n,\pm}(t)| \leq 2M \quad (33b)$$

The dependence of κ on h_n , as prescribed in (29), implies:

$$C := \liminf_{n \rightarrow \infty} h_n R_2^{-1}(R(2M)) > 0 \quad (33c)$$

so that, by (33b) for large enough n

$$R_2^{-1}\left(R(D^+ \psi_j^{n,\pm}(t))\right) > \frac{C}{2h_n} \text{ for all } j \notin S\left(\psi_j^{n,\pm}(0)\right) \text{ and } t \in [0, T]. \quad (33d)$$

Choose $\delta < C/4$. For large enough n , we will certainly have:

$$(-1)^{\alpha(j)} \psi_j^{n,-}(t) < u_j^{h_n}(t) < (-1)^{\alpha(j)} \psi_j^{n,+}(t) \quad (33e)$$

for some positive time. Then (32a), (33a), and (33e) verify hypothesis 1 and 3 of Proposition 2 for positive time. Meanwhile, (33d), (33e), and the choice of δ verify hypothesis 2 of Proposition 2 at $t = 0$. But as in the proof of Theorem 1, since (33d) holds for all $t \in [0, T]$, these are sufficient to ensure that the three hypothesis of Proposition 2 are satisfied for all $t \in [0, T]$ and lead to the conclusion that we want:

$$(-1)^{\alpha(j)} \psi_j^{n,-}(t) \leq u_j^{h_n}(t) \leq (-1)^{\alpha(j)} \psi_j^{n,+}(t) \text{ for all } t \in [0, T]. \quad (34)$$

At this stage, it is possible to repeat the proof of Corollary 1 to get an estimate on how fast jumps of height less than 2δ vanish; the upper bound one obtains goes to 0 when we send first $n \rightarrow \infty$ and then $\delta \rightarrow 0$. \square

Remark: In some sense, Theorem 2 says that given an initial piecewise smooth signal corrupted by noise (i.e. a signal with a few large and many small jumps), the Perona-Malik scheme de-noises the signal by quickly removing the small jumps and maintaining the large ones. This stability property hinged on inequality (33c)

in our proof, and is a result of the choice of the constitutive function (26) and scaling (29). It is in contrast to the different function and scaling considered in [3] that lead to a different continuum limit for which such a stability property is not to be expected.

Acknowledgement: The author would like to thank his former advisor Robert V. Kohn for his continuing attention and encouragement.

References

- [1] Chambolle, A. Image segmentation by variational methods: Mumford and Shah functional and discrete approximations. *SIAM J. Appl. Math.* **55** (1995), no. 3, 827-863.
- [2] Esedoglu, S. An Analysis of the Perona-Malik Scheme. *Comm. Pure Appl. Math.* **54** (2001), 1442-1487.
- [3] Gobbino, M. Gradient flow for the one-dimensional Mumford-Shah functional. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **27** (1998), no. 1, 145-193 (1999).
- [4] Kawohl, B.; Kutev, N. Maximum and comparison principle for one-dimensional anisotropic diffusion. *Math. Ann.* **311** (1998), no. 1, 107-123.
- [5] Kichenassamy, S. The Perona-Malik paradox. *SIAM J. Appl. Math.* **57** (1997), no. 5, 1328-1342.
- [6] Perona, P.; Malik, J. Scale space and edge detection using anisotropic diffusion. *Proc. IEEE Computer Soc. Workshop on Computer Vision.* (1987), 16-22.
- [7] Perona, P.; Malik, J. Scale space and edge detection using anisotropic diffusion. Dept. of EECS Technical Report, U.C. Berkeley, 1988.
- [8] Perona, P.; Shiota, T.; Malik, J. Anisotropic diffusion. *Geometry-Driven Diffusion in Computer Vision*, 73-92. Edited by B. ter Haar Romeny. Computational Imaging and Vision, 1. Kluwer Academic Publishers, Dordrecht, 1994.
- [9] Yuo, Y.; Xu, W.; Tannenbaum, A.; Kaveh, M. Behavioral analysis of anisotropic diffusion in image processing. *IEEE Trans. Image Process.* **5** (1996), no. 11, 1539-1553.

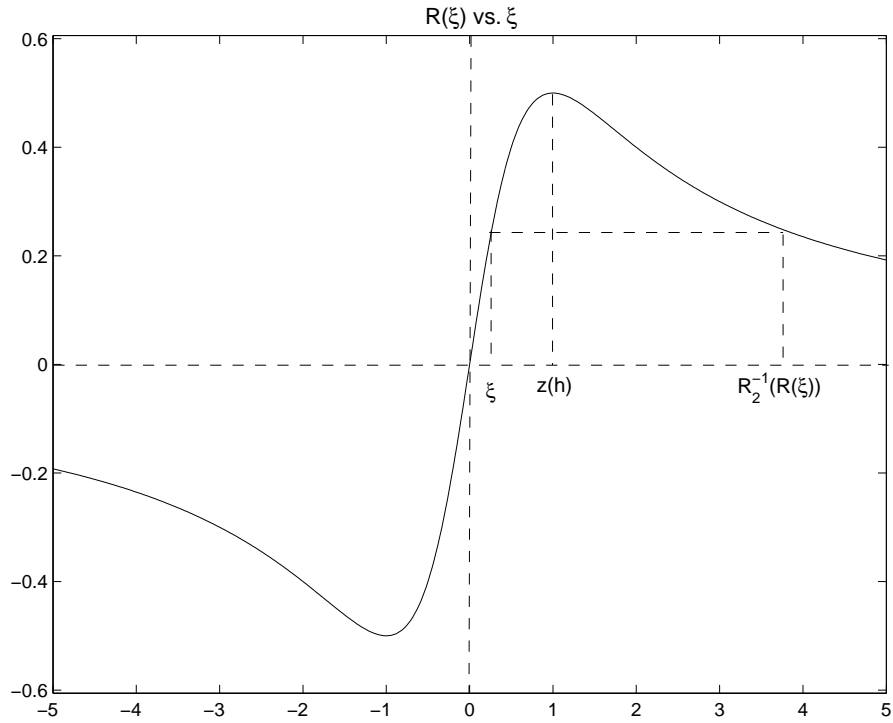


Figure 1: A typical choice for the function $R_h(\xi)$ that appears in the scheme. Here $R_h(\xi) = \xi/(1 + \xi^2/h)$ with $h = 1$.

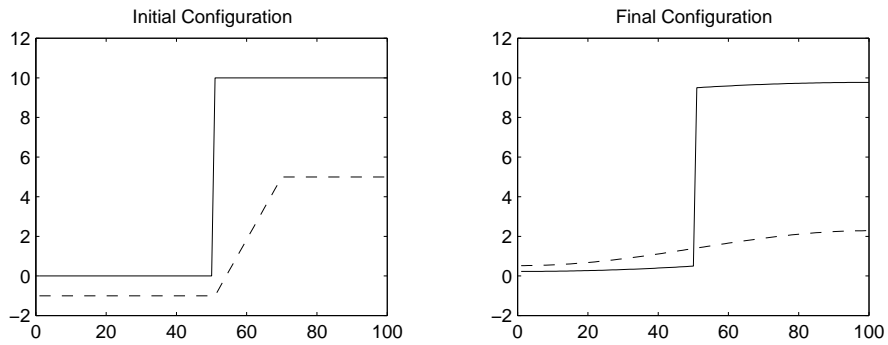


Figure 2: Violation of comparison principle. Solution represented by the dashed line remains in the well-posed (parabolic) regime at all times. Nevertheless, order is lost. It is very easy to understand why this happens.

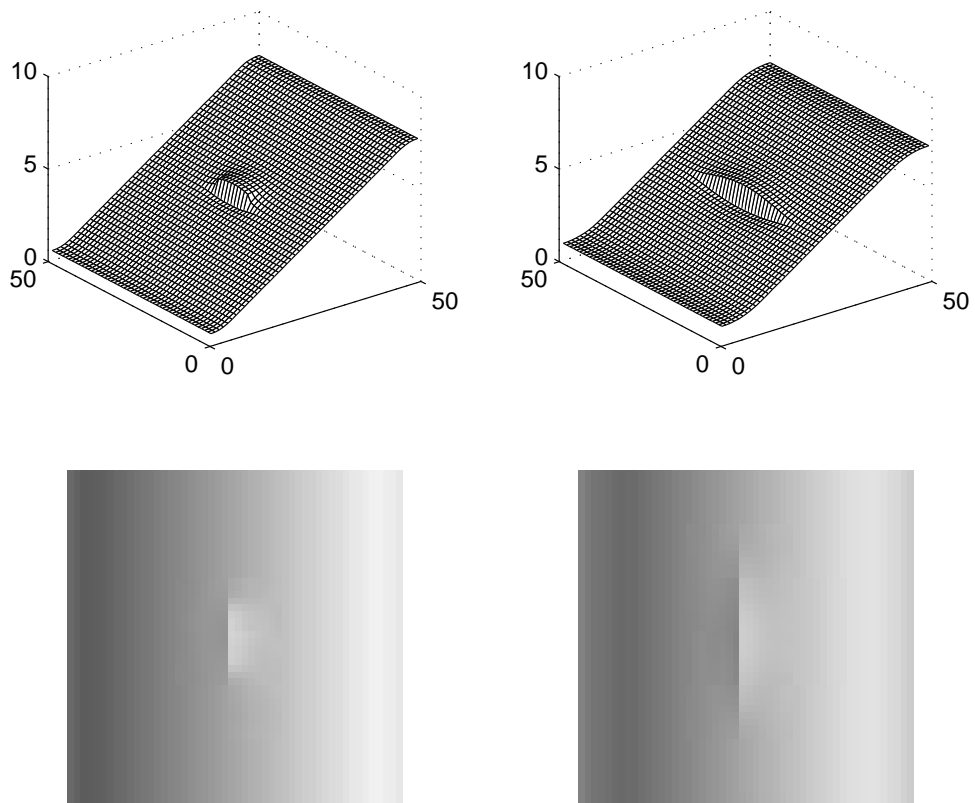


Figure 3: Jump set (or the supersonic regime) can grow in two-dimensions. Here, a small “crack” in the initial data propagates.

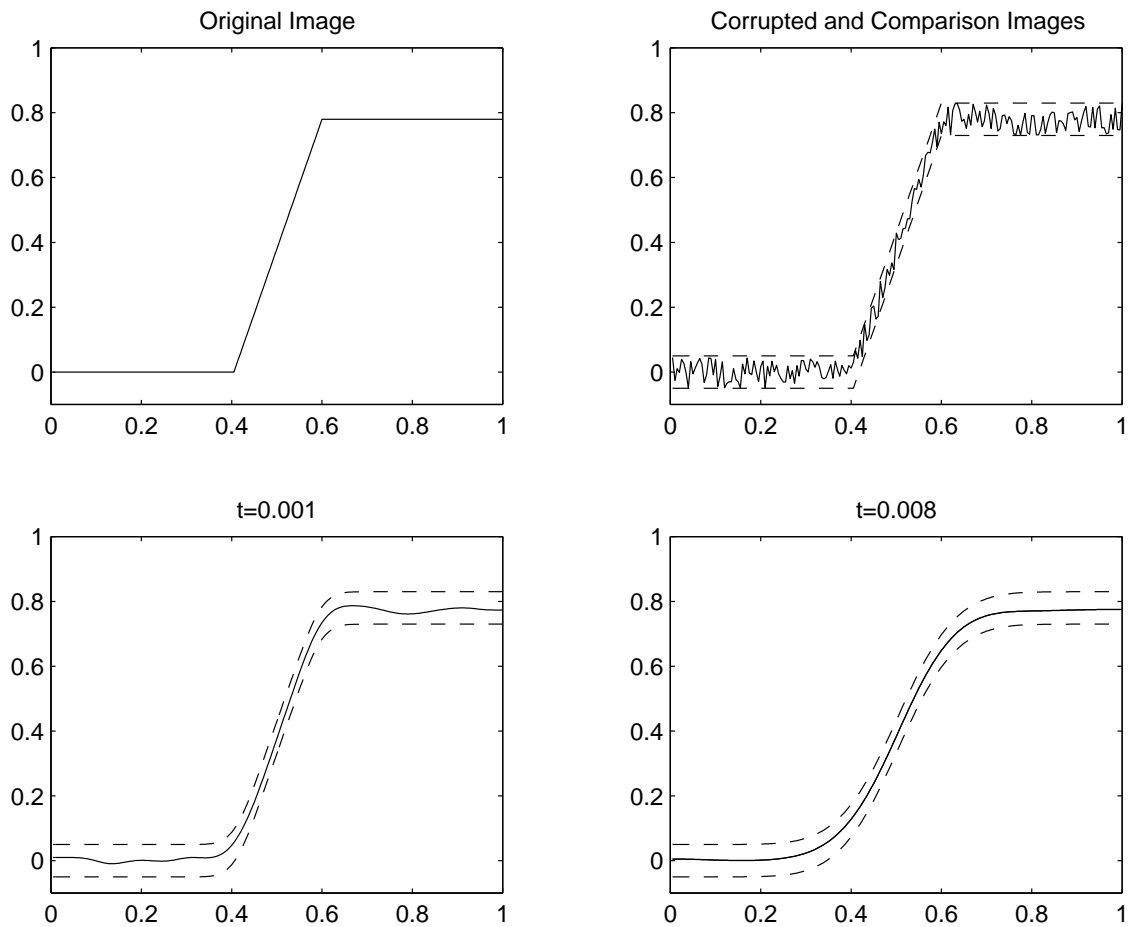


Figure 4: An application of the comparison principle. The corrupted initial data (which is not subsonic) quickly becomes subsonic, and no artificial “edges” are introduced. The noise amplitude here is 0.05, and the maximum amplitude allowed by Theorem 1 is 0.0525.

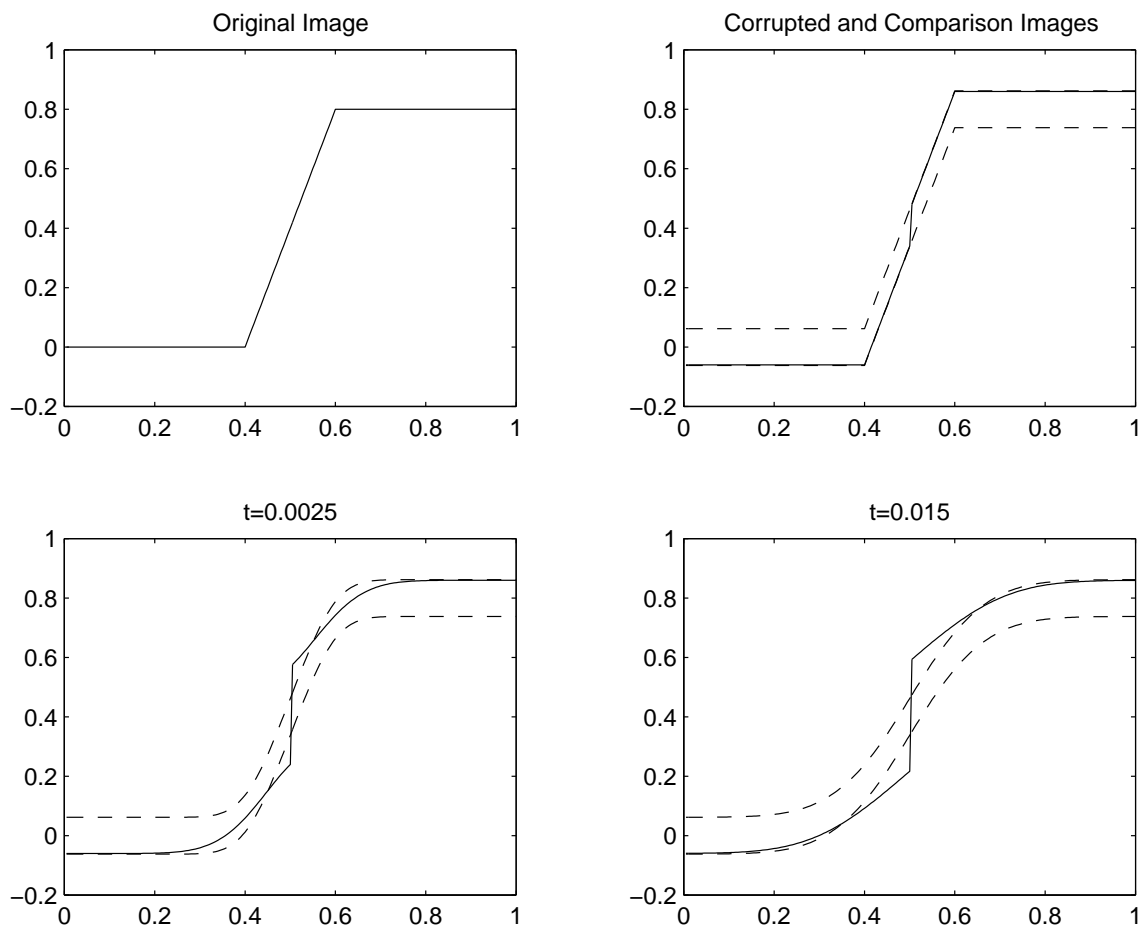


Figure 5: Effects of a perturbation with a (deliberately chosen) noise of amplitude that exceeds the maximum amplitude allowed. Here, the noise amplitude is 0.06, higher than what Theorem 1 allows, which is 0.0525. The result is an artificial “edge”.

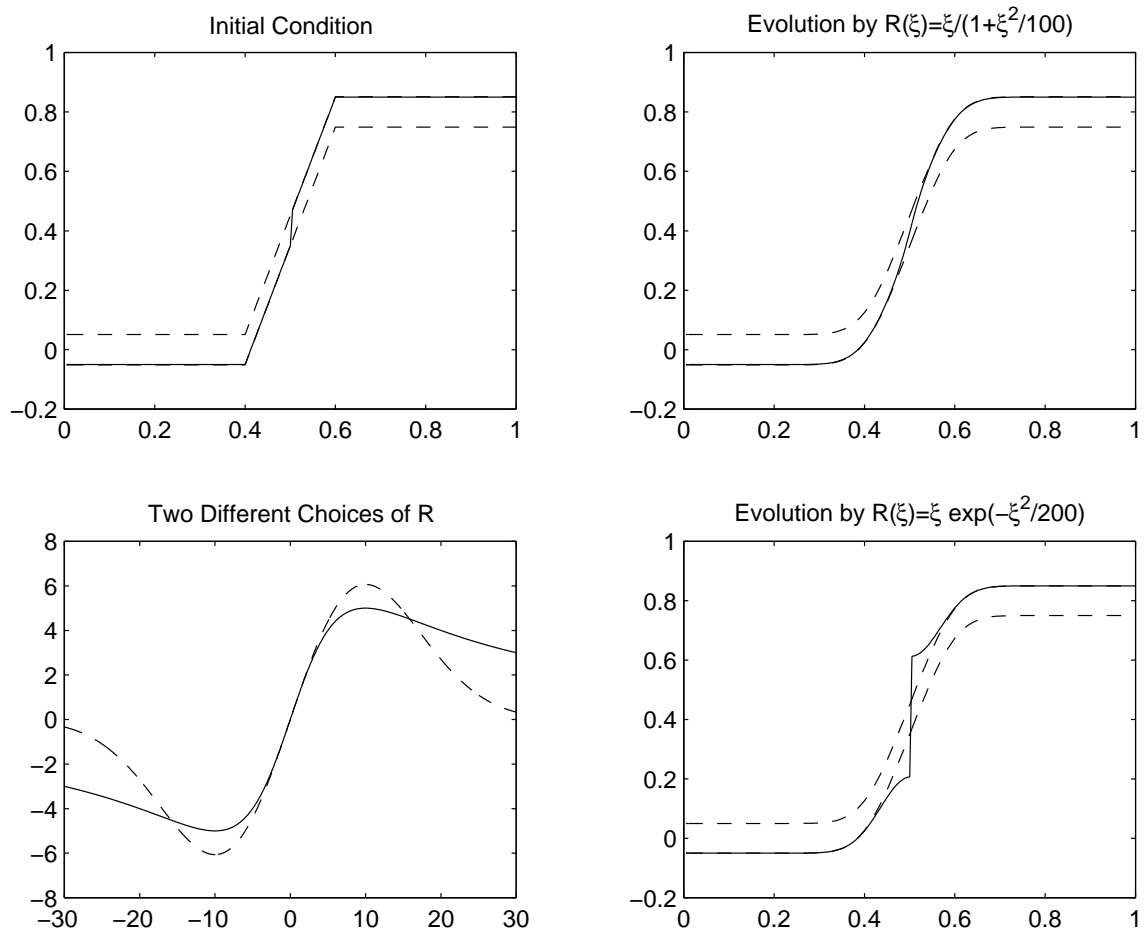


Figure 6: Evolution via different choices for the function R_{κ} . The upper left hand figure illustrates the initial signal and comparison functions; the initial signal was obtained from that of Example 1 (Figure 4) after modifying by a deliberately chosen perturbation. The lower left hand figure shows the two distinct choices for the function $R(\xi)$ used in this example. The second column shows the evolution of the initial image and comparison functions with $R(\xi) = \xi/(1 + \xi^2/100)$ and $R(\xi) = \xi \exp(-\xi^2/200)$. For both of these functions, $z(\kappa) = 10$.