A Subsonic-Well-Balanced Reconstruction Scheme for Shallow Water Flows

François Bouchut, Tomás Morales De Luna

A SUBSONIC-WELL-BALANCED RECONSTRUCTION SCHEME FOR SHALLOW WATER FLOWS

FRANÇOIS BOUCHUT† AND TOMÁS MORALES DE LUNA‡

Abstract. We consider the Saint-Venant system for shallow water flows with nonflat bottom. In past years, efficient well-balanced methods have been proposed in order to well resolve solutions close to steady states at rest. Here we describe a strategy based on a local subsonic steady state reconstruction that allows one to derive a subsonic-well-balanced scheme, preserving exactly all the subsonic steady states. It generalizes the now well-known hydrostatic solver, and like the latter it preserves the nonnegativity of the water height and satisfies a semidiscrete entropy inequality. An application to the Euler–Poisson system is proposed.

Key words. shallow water, subsonic reconstruction, subsonic steady states, well-balanced scheme, semidiscrete entropy inequality

1. Introduction. We consider the classical Saint-Venant system for shallow water flows with topography. It is a hyperbolic system of conservation laws that approximately describes various geophysical flows, such as rivers, coastal areas, oceans when completed with a Coriolis term, and granular flows when completed with friction. Numerical approximate solutions to this system can be generated using conservative finite volume methods, which are known to properly handle shocks and contact discontinuities. As is now well known, in the occurrence of source terms such as topography, a classical centered discretization does not allow precise computations for near steady states. One then has to use the so-called well-balanced schemes, which properly balance the fluxes and the source at the level of each interface. Such schemes have been proposed in [13], [14], [4], [21], [12], [20], [15], [8], [16], [3], [5], [11], [10], [2], [6], [1], [9], [17].

Additionally to the well-balanced property, the difficulty is to also have schemes that satisfy very natural properties such as the conservativity of the water height ρ, the nonnegativity of ρ, the ability to compute dry states $\rho = 0$ and transcritical flows when the Jacobian matrix F' of the flux function becomes singular, and eventually the ability to satisfy a discrete entropy inequality. The solvers satisfying all these requirements are very few; they are those obtained by exact resolution in [10], by the kinetic method of [20], by the hydrostatic reconstruction method of [1], and by the Suliciu relaxation method of [7].

Nevertheless, these solvers usually preserve only the steady states at rest, for which $u \equiv 0$. Some schemes that are able to maintain all steady states have been proposed in [10], [9], [18], [19] but do not satisfy all of the desired properties cited

†Laboratoire d'Analyse et de Mathèmatiques Appliquées, LAMA - UMR 8050, Université Paris-Est - Marne-la-Vallée, 5 boulevard Descartes, Cité Descartes - Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France (francois.bouchut@univ-mlv.fr).

‡Corresponding author. Departamento de Matemáticas, Edificio Albert Einstein (C2), Campus de Rabanales, Universidad de Córdoba, 14071 Córdoba, Spain (Tomas.Morales@uco.es).

before (entropy satisfying and the conservativity and nonnegativity of ρ). The object of this paper is to go further in well-balanced schemes by building a solver with all the above requirements and, overall, the property of preserving exactly all subsonic steady states.

2. Saint-Venant system and well-balanced schemes. The Saint-Venant system describes the evolution of the water height $\rho(t,x)$ and the velocity $u(t,x)$ in the horizontal direction, of a thin layer of water flowing over a slowly varying topography. In one space dimension, the system writes as

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) + \rho g z_x = 0, \end{cases} \tag{2.1}$$

where $g > 0$ is the gravitational constant and $z(x)$ is the topography. We shall denote

$$Z = gz. \tag{2.2}$$

The physically relevant case is $p(\rho) = g\rho^2/2$, but we shall deal with the general case $p(\rho)$. We shall assume as usual that $p' > 0$, and we suppose that

$$\rho^2 p'(\rho) \text{ is strictly increasing,} \qquad \rho^2 p'(\rho) \to \infty \text{ as } \rho \to \infty, \tag{2.3}$$

$$\int_0^1 \frac{p'(\rho)}{\rho}\, d\rho < \infty, \qquad p'(\rho) \to 0 \text{ as } \rho \to 0, \qquad \int_1^\infty \frac{p'(\rho)}{\rho}\, d\rho = \infty. \tag{2.4}$$

These assumptions are satisfied in particular for the pressure law of isentropic gas dynamics $p(\rho) = \kappa\rho^\gamma$, with $\gamma > 1$ and $\kappa > 0$. We define as usual the internal energy $e(\rho)$ by

$$e'(\rho) = \frac{p(\rho)}{\rho^2}. \tag{2.5}$$

Note that the integrability conditions in (2.4) imply that $e(\rho) + p(\rho)/\rho$ has a finite limit as $\rho \to 0$ and tends to ∞ as $\rho \to \infty$. For future reference we denote the flux by

$$F(U) = \Big(\rho u, \rho u^2 + p(\rho)\Big), \qquad U = (\rho, \rho u). \tag{2.6}$$

The Saint-Venant model is very robust, being hyperbolic and admitting an entropy inequality related to the physical energy,

$$\partial_t\, \widetilde{\eta}(U,Z) + \partial_x\, \widetilde{G}(U,Z) \le 0, \tag{2.7}$$

where

$$\begin{aligned} &\eta(U) = \rho u^2/2 + \rho e(\rho), && G(U) = \big(\rho u^2/2 + \rho e(\rho) + p(\rho)\big)\, u, \\ &\widetilde{\eta}(U,Z) = \eta(U) + \rho Z, && \widetilde{G}(U,Z) = G(U) + \rho u Z. \end{aligned} \tag{2.8}$$

The steady states of (2.1) can be described as follows. We subtract u times the first equation in (2.1) from the second and divide the result by ρ. We get

$$\partial_t u + \partial_x \left(\frac{u^2}{2} + e(\rho) + \frac{p(\rho)}{\rho} + Z \right) = 0. \tag{2.9}$$

Therefore, the steady states are exactly the functions $\rho(x)$, $u(x)$ satisfying

$$\begin{cases} \rho u = Cst, \\ \frac{u^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho) + Z = Cst. \end{cases} \tag{2.10}$$

In particular, we have the so-called steady state at rest

$$u = 0, \qquad e + \frac{p}{\rho} + Z = Cst. \tag{2.11}$$

As exposed in [7], a first-order finite volume method for solving (2.1) writes generically with $U = (\rho, \rho u)$ as

$$U_i^{n+1} - U_i + \frac{\Delta t}{\Delta x_i}\left(F_{i+1/2-} - F_{i-1/2+}\right) = 0, \tag{2.12}$$

$$F_{i+1/2-} = \mathcal{F}_l(U_i, U_{i+1}, \Delta Z_{i+1/2}), \qquad F_{i+1/2+} = \mathcal{F}_r(U_i, U_{i+1}, \Delta Z_{i+1/2}) \tag{2.13}$$

for some left/right numerical fluxes $\mathcal{F}_l(U_l, U_r, \Delta Z)$, $\mathcal{F}_r(U_l, U_r, \Delta Z)$, with

$$\Delta Z_{i+1/2} = Z_{i+1} - Z_i, \tag{2.14}$$

and where Δx_i denotes a possibly variable mesh size, $\Delta x_i = x_{i+1/2} - x_{i-1/2}$. We then have the following characterizations (see [7]).

▷ The conservativity or density writes, with $\mathcal{F}_{l/r} = (\mathcal{F}^{\rho}_{l/r}, \mathcal{F}^{\rho u}_{l/r})$,

$$\mathcal{F}^{\rho}_l = \mathcal{F}^{\rho}_r \equiv \mathcal{F}^{\rho}. \tag{2.15}$$

▷ The consistency-conservativity can be written as

$$\begin{cases} \mathcal{F}^{\rho}(U, U, 0) = \rho u, \\ \mathcal{F}^{\rho u}_l(U, U, 0) = \mathcal{F}^{\rho u}_r(U, U, 0) = \rho u^2 + p(\rho), \end{cases} \tag{2.16}$$

(2.17)
$$\mathcal{F}^{\rho u}_r(U_l, U_r, \Delta Z) - \mathcal{F}^{\rho u}_l(U_l, U_r, \Delta Z) = -\rho \Delta Z + o(\Delta Z) \quad \text{as } U_l, U_r \to U,\ \Delta Z \to 0.$$

▷ The well-balancing property can be stated as the property of having, for the considered steady states,

$$\mathcal{F}_l(U_l, U_r, \Delta Z) = F(U_l), \qquad \mathcal{F}_r(U_l, U_r, \Delta Z) = F(U_r). \tag{2.18}$$

▷ The property of satisfying a semidiscrete entropy inequality (i.e., related to the limit $\Delta t \to 0$) is characterized by the existence of a numerical entropy flux $\widetilde{\mathcal{G}}(U_l, U_r, Z_l, Z_r)$ consistent with the exact flux $\widetilde{G}(U, Z)$, such that

$$\widetilde{G}(U_r, Z_r) + \widetilde{\eta}'(U_r, Z_r)(\mathcal{F}_r(U_l, U_r, \Delta Z) - F(U_r)) \le \widetilde{\mathcal{G}}(U_l, U_r, Z_l, Z_r), \tag{2.19}$$

$$\widetilde{\mathcal{G}}(U_l, U_r, Z_l, Z_r) \le \widetilde{G}(U_l, Z_l) + \widetilde{\eta}'(U_l, Z_l)(\mathcal{F}_l(U_l, U_r, \Delta Z) - F(U_l)), \tag{2.20}$$

where $\widetilde{\eta}'(U, Z)$ is the derivative of $\widetilde{\eta}(U, Z)$ with respect to U.

The hydrostatic reconstruction scheme satisfies all the above and is defined as

$$
\begin{aligned}
\mathcal{F}_l(U_l, U_r, \Delta Z) &= \mathcal{F}(U_l^*, U_r^*) + \begin{pmatrix} 0 \\ p(\rho_l) - p(\rho_l^*) \end{pmatrix}, \\
\mathcal{F}_r(U_l, U_r, \Delta Z) &= \mathcal{F}(U_l^*, U_r^*) + \begin{pmatrix} 0 \\ p(\rho_r) - p(\rho_r^*) \end{pmatrix},
\end{aligned} \tag{2.21}
$$

where $\mathcal{F}(U_l, U_r)$ is a numerical flux for the shallow water problem without source ($Z = cst$), and the reconstructed states U_l^*, U_r^* are defined by

$$
U_l^* = (\rho_l^*, \rho_l^* u_l), \qquad U_r^* = (\rho_r^*, \rho_r^* u_r), \tag{2.22}
$$

$$
\begin{aligned}
\Big(e + p/\rho\Big)(\rho_l^*) &= \Big((e + p/\rho)(\rho_l) - (\Delta Z)_+\Big)_+, \\
\Big(e + p/\rho\Big)(\rho_r^*) &= \Big((e + p/\rho)(\rho_r) - (-\Delta Z)_+\Big)_+,
\end{aligned} \tag{2.23}
$$

where we use the notation $X_+ = \max(0, X)$, and we assume that $e(\rho) + p(\rho)/\rho \to 0$ as $\rho \to 0$.

3. Well-balanced scheme with subsonic reconstruction. We would now like to explain how it is possible to extend the hydrostatic reconstruction scheme (2.21)–(2.23) in order to obtain a scheme that satisfies the above requirements and preserves some more general steady states than the steady states at rest. We shall obtain in particular a scheme that preserves all subsonic steady states, that is, the steady states that verify $u^2 < p'(\rho)$. This property will be called subsonic-well-balanced. Note in particular that the steady states at rest (with $u = 0$) are subsonic.

3.1. Parametrization of numerical fluxes. Following [1], [7], we propose and analyze finite volume schemes defined by (2.12), (2.13) with numerical fluxes

$$
\begin{aligned}
\mathcal{F}_l(U_l, U_r, \Delta Z) &= \mathcal{F}(U_l^*, U_r^*) + \begin{pmatrix} 0 \\ p(\rho_l) - p(\rho_l^*) + T_l(U_l, U_r, \Delta Z) \end{pmatrix}, \\
\mathcal{F}_r(U_l, U_r, \Delta Z) &= \mathcal{F}(U_l^*, U_r^*) + \begin{pmatrix} 0 \\ p(\rho_r) - p(\rho_r^*) + T_r(U_l, U_r, \Delta Z) \end{pmatrix},
\end{aligned} \tag{3.1}
$$

where $\mathcal{F}$ stands for a numerical flux for the homogeneous problem ($Z = cst$), and the interface values U_l^*, U_r^* are derived from a local reconstruction procedure. They should satisfy at least that $U_l^* = U_l$, $U_r^* = U_r$ when $\Delta Z = 0$.

The extra terms T_l, T_r appear here in order to balance the advection term $\partial_x(\rho u^2)$ in (2.1), which was not considered in the hydrostatic scheme. Taking into account the constant discharge condition in (2.10), this balancing requirement suggests the relations

$$
T_l = \rho_l u_l (u_l - u_l^*), \qquad T_r = \rho_r u_r (u_r - u_r^*). \tag{3.2}
$$

It is obvious from the characterization (2.18) that a steady state is maintained exactly by the scheme (2.12), (2.13), (3.1) if, for such a state, the reconstructed states satisfy $U_l^* = U_r^*$, $\rho_l^* u_l^* = \rho_l u_l$, and $\rho_r^* u_r^* = \rho_r u_r$, and (3.2) is satisfied.

Taking into account the steady state equation (2.10), one could guess a reconstruction of the states U_l^*, U_r^* to be

$$
\left\{
\begin{aligned}
&\frac{(u_l^*)^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_l^*) + Z^* = \frac{u_l^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_l) + Z_l, \\
&\rho_l^* u_l^* = \rho_l u_l,
\end{aligned}
\right. \tag{3.3}
$$

$$
(3.4)\qquad \begin{cases} \dfrac{(u_r^*)^2}{2} + \left(e + \dfrac{p}{\rho}\right)(\rho_r^*) + Z^* = \dfrac{u_r^2}{2} + \left(e + \dfrac{p}{\rho}\right)(\rho_r) + Z_r, \\ \rho_r^* u_r^* = \rho_r u_r, \end{cases}
$$

with

$$
(3.5)\qquad Z^* = \max(Z_l, Z_r).
$$

There exist solutions to the previous system if ΔZ is small enough (recall that $\Delta Z = Z_r - Z_l$), but this is not true for arbitrary ΔZ, as we shall see later. The idea is thus to consider generalized T_l, T_r, to be defined later. Their definition is motivated by the entropy inequality.

LEMMA 3.1. *Let $\mathcal{F}(U_l, U_r)$ be a given consistent numerical flux for the Saint-Venant problem without source that verifies a semidiscrete entropy inequality for the entropy pair (η, G) given by (2.8), and denote $\mathcal{F} = (\mathcal{F}^\rho, \mathcal{F}^{\rho u})$. A sufficient condition for the scheme (2.12), (2.13), (3.1) to be semidiscrete entropy satisfying for the entropy pair $(\widetilde{\eta}, \widetilde{G})$ in (2.8) is that for some Z^**

$$
(3.6)\qquad \begin{aligned} &G(U_l^*) + \eta'(U_l^*)(\mathcal{F}(U_l^*, U_r^*) - F(U_l^*)) + \mathcal{F}^\rho(U_l^*, U_r^*)Z^* \\ &\leq G(U_l) + \eta'(U_l)(\mathcal{F}_l(U_l, U_r, \Delta Z) - F(U_l)) + \mathcal{F}^\rho(U_l^*, U_r^*)Z_l \end{aligned}
$$

and

$$
(3.7)\qquad \begin{aligned} &G(U_r) + \eta'(U_r)(\mathcal{F}_r(U_l, U_r, \Delta Z) - F(U_r)) + \mathcal{F}^\rho(U_l^*, U_r^*)Z_r \\ &\leq G(U_r^*) + \eta'(U_r^*)(\mathcal{F}(U_l^*, U_r^*) - F(U_r^*)) + \mathcal{F}^\rho(U_l^*, U_r^*)Z^*. \end{aligned}
$$

Proof. The numerical flux $\mathcal{F}$ satisfies a semidiscrete entropy inequality associated with the entropy pair (η, G); thus

$$
(3.8)\qquad \begin{aligned} &G(U_r) + \eta'(U_r)(\mathcal{F}(U_l, U_r) - F(U_r)) \\ &\leq \mathcal{G}(U_l, U_r) \leq G(U_l) + \eta'(U_l)(\mathcal{F}(U_l, U_r) - F(U_l)) \end{aligned}
$$

for a numerical flux $\mathcal{G}$ consistent with G. For the scheme (2.12), (2.13), (3.1) to be semidiscrete entropy satisfying for the entropy pair $(\widetilde{\eta}, \widetilde{G})$, (2.19), (2.20) should hold. Let

$$
(3.9)\qquad \widetilde{\mathcal{G}}(U_l, U_r, Z_l, Z_r) = \mathcal{G}(U_l^*, U_r^*) + \mathcal{F}^\rho(U_l^*, U_r^*)Z^*.
$$

As $\mathcal{G}$ is consistent with G, $\widetilde{\mathcal{G}}$ is consistent with $\widetilde{G}$. The comparison between (3.8) evaluated at (U_l^*, U_r^*) and (2.19) and (2.20) gives that (3.6) and (3.7) are sufficient conditions. ∎

LEMMA 3.2. *Denote $(\mathcal{F}^\rho, \mathcal{F}^{\rho u}) \equiv \mathcal{F}(U_l^*, U_r^*)$, $T_l \equiv T_l(U_l, U_r, \Delta Z)$, and $T_r \equiv T_r(U_l, U_r, \Delta Z)$, and define the quantities*

$$
(3.10)\qquad \begin{aligned} W_l \equiv \mathcal{F}^\rho \cdot &\left(\left(e + \frac{p}{\rho}\right)(\rho_l) - \left(e + \frac{p}{\rho}\right)(\rho_l^*) + Z_l - Z^* + \frac{(u_l^*)^2}{2} - \frac{u_l^2}{2}\right) \\ &+ (u_l - u_l^*)\left(\mathcal{F}^{\rho u} - p(\rho_l^*)\right) + u_l T_l, \end{aligned}
$$

$$
(3.11)\qquad \begin{aligned} W_r \equiv \mathcal{F}^\rho \cdot &\left(\left(e + \frac{p}{\rho}\right)(\rho_r) - \left(e + \frac{p}{\rho}\right)(\rho_r^*) + Z_r - Z^* + \frac{(u_r^*)^2}{2} - \frac{u_r^2}{2}\right) \\ &+ (u_r - u_r^*)\left(\mathcal{F}^{\rho u} - p(\rho_r^*)\right) + u_r T_r. \end{aligned}
$$

A necessary and sufficient condition for (3.6), (3.7) *to hold is that*

$$W_l \geq 0, \qquad W_r \leq 0. \tag{3.12}$$

Proof. From the explicit value of F, G and computing $\eta'(U) = (e(\rho) + p(\rho)/\rho - u^2/2, u)$, one gets the identity $G(U) - \eta'(U)F(U) = -u\,p(\rho)$. Plugging this into (3.6), (3.7) yields the result. ◻

We remark that in the particular case when U_l^*, U_r^* satisfy (3.3)–(3.4), we have

$$\begin{aligned} W_l &= (u_l^* - u_l)\Big((u_l + u_l^*)\mathcal{F}^\rho - \mathcal{F}^{\rho u} + p(\rho_l^*)\Big) + u_l T_l, \\ W_r &= (u_r^* - u_r)\Big((u_r + u_r^*)\mathcal{F}^\rho - \mathcal{F}^{\rho u} + p(\rho_r^*)\Big) + u_r T_r. \end{aligned} \tag{3.13}$$

Thus, the choice of T_l, T_r given by (3.2) makes $W_l = W_r = 0$ whenever $U_l^* = U_r^*$.

3.2. Subsonic reconstruction. We intend here to define reconstructed states that satisfy (3.3), (3.4). Nevertheless, it will not be always possible to do so, as we will see in what follows. The possibility of achieving these relations is related to the following definitions.

Definition 3.3. *Let $\rho \geq 0$, and let $u \in \mathbb{R}$. We say that (ρ, u) is a sonic, subsonic, or supersonic point for the Saint-Venant system* (2.1) *if we have, respectively, $u^2 = p'(\rho)$, $u^2 < p'(\rho)$, or $u^2 > p'(\rho)$.*

Definition 3.4. *Let $q \in \mathbb{R}$. We define $\rho_s(q)$ as the solution to*

$$\rho_s^2 p'(\rho_s) = q^2, \quad \rho_s \geq 0, \tag{3.14}$$

and

$$m_s(q) = \left(e + \frac{p}{\rho} + \frac{p'}{2}\right)(\rho_s(q)). \tag{3.15}$$

Because of the assumptions (2.3), (2.4) on p, there exists a unique ρ_s solution to (3.14). Moreover, $m_s(0)$ is well defined:

$$m_s(0) = \left(e + \frac{p}{\rho}\right)(0). \tag{3.16}$$

In the particular case when $p(\rho) = \kappa\rho^\gamma$, we have

$$\rho_s(q) = \left(\frac{q^2}{\kappa\gamma}\right)^{\frac{1}{\gamma+1}}, \quad m_s(q) = \left(\frac{1}{2} + \frac{1}{\gamma - 1}\right)(\kappa\gamma)^{\frac{2}{\gamma+1}} |q|^{2\frac{\gamma-1}{\gamma+1}}. \tag{3.17}$$

The following proposition gives the necessary and sufficient conditions for the existence of a solution to (3.3) and (3.4).

Proposition 3.5. *Let $u_0 \in \mathbb{R}, \rho_0 \geq 0, \delta \in \mathbb{R}$. Consider the system*

$$\begin{cases} \frac{(u^*)^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho^*) = \frac{u_0^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_0) + \delta, \\ \rho^* u^* = \rho_0 u_0, \\ \rho^* \geq 0, \ u^* \in \mathbb{R}, \end{cases} \tag{3.18}$$

and denote $\rho_s \equiv \rho_s(\rho_0 u_0)$. There exists a solution (ρ^*, u^*) to (3.18) *if and only if*

$$\frac{u_0^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_0) + \delta \geq m_s(\rho_0 u_0). \tag{3.19}$$

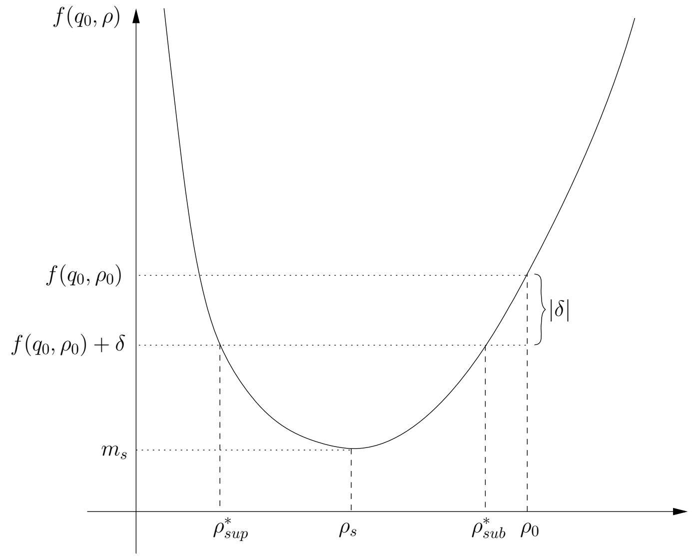

FIG. 1. *Function* $f(q_0,\cdot)$.

Moreover, the following hold:

(i) *If we have equality in* (3.19)*, there is only one solution* (ρ^*,u^*) *to* (3.18)*, and it is given by*

$$\rho^* = \rho_s, \qquad u^* = \begin{cases} \frac{\rho_0 u_0}{\rho_s} & \text{if } \rho_0 u_0 \neq 0, \\ 0 & \text{if } \rho_0 u_0 = 0. \end{cases} \tag{3.20}$$

(ii) *If we have a strict inequality in* (3.19)*, then there are exactly two different solutions* $(\rho^*_{sup}, u^*_{sup})$ *and* $(\rho^*_{sub}, u^*_{sub})$ *to* (3.18)*, with* $\rho^*_{sup} \le \rho_s < \rho^*_{sub}$*, and* $\rho^*_{sup} < \rho_s$ *for* $\rho_0 u_0 \neq 0$.

(iii) *A solution* (ρ^*,u^*) *to* (3.18)*, with* $\rho^* u^* \neq 0$*, is a sonic (resp., subsonic or supersonic) point if and only if* $\rho^* = \rho_s$ *(resp.,* $\rho^* > \rho_s$ *or* $\rho^* < \rho_s$*).*

Proof. Let us suppose first that $\rho_0 u_0 \neq 0$, and consider the function

$$\begin{aligned} f : \mathbb{R} \times (0,\infty) &\longrightarrow \mathbb{R}, \\ (q,\rho) \qquad &\mapsto f(q,\rho) = \frac{q^2}{2\rho^2} + \left(e + \frac{p}{\rho}\right)(\rho). \end{aligned} \tag{3.21}$$

Then (ρ^*,u^*) is a solution to (3.18) if and only if $\rho^* u^* = \rho_0 u_0$ and $f(\rho_0 u_0, \rho^*) = f(\rho_0 u_0, \rho_0) + \delta$. We have $\frac{\partial f}{\partial \rho}(q,\rho) = (\rho^2 p'(\rho) - q^2)/\rho^3$; thus, according to (2.3), (2.4), f is strictly increasing in $(\rho_s(q),\infty)$ and strictly decreasing in $(0,\rho_s(q))$. Therefore, $\rho_s(q)$ is a minimum point of f with minimum value $f(q,\rho_s(q)) = m_s(q)$. Figure 1 shows a sketch of the function $f(\rho_0 u_0,\cdot)$ and the solutions ρ^*_{sub}, ρ^*_{sup} in the case $\delta < 0$. Thus, condition (3.19) follows, as well as (i) and (ii).

Now, if we consider (ρ^*,u^*) a solution of (3.18), we have

$$(u^*)^2 > p'(\rho^*) \Leftrightarrow (\rho_0 u_0)^2 > (\rho^*)^2 p'(\rho^*). \tag{3.22}$$

Since $\rho^2 p'(\rho)$ is a strictly increasing function with $\rho_s^2 p'(\rho_s) = (\rho_0 u_0)^2$,

$$(u^*)^2 > p'(\rho^*) \Leftrightarrow \rho^* < \rho_s, \tag{3.23}$$

which proves (iii).

In the case $\rho_0 u_0 = 0$, the second condition in (3.18) simplifies to $\rho^* u^* = 0$. Thus the system can have solutions with either $\rho^* = 0$, giving $(u^*)^2/2 = \text{r.h.s.} - m_s(0)$, or

$u^* = 0$, giving $(e + p/\rho)(\rho^*) = \text{r.h.s.}$ One easily sees that a solution exists if and only if r.h.s. $\geq m_s(0)$, proving condition (3.19). In case of equality, the only solution is $\rho^* = 0$, $u^* = 0$, which is sonic. In case of inequality, the solutions are given by

$$\rho^*_{sup} = 0, \quad \frac{(u^*_{sup})^2}{2} = \frac{u_0^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_0) + \delta - m_s(0), \tag{3.24}$$

$$u^*_{sub} = 0, \quad \left(e + \frac{p}{\rho}\right)(\rho^*_{sub}) = \frac{u_0^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_0) + \delta. \tag{3.25}$$

This gives the result, with the convention that for (3.24) we identify the two solutions having density $\rho^*_{sup} = 0$. Note that these solutions are supersonic, while the ones corresponding to (3.25) are subsonic. □

COROLLARY 3.6. *Let $\rho_0 > 0$, and let $u_0 \in \mathbb{R}\backslash\{0\}$. Then (ρ_0, u_0) is a sonic (resp., subsonic or supersonic) point if and only if $\rho_0 = \rho_s(\rho_0 u_0)$ (resp., $\rho_0 > \rho_s(\rho_0 u_0)$ or $\rho_0 < \rho_s(\rho_0 u_0)$).*

Proof. The point (ρ_0, u_0) is a trivial solution to (3.18) with $\delta = 0$. Thus the result follows from (iii) in the previous proposition. □

LEMMA 3.7. *Suppose that in Proposition 3.5 we are in the case of two solutions ρ^*_{sub}, ρ^*_{sup}. Then one has the following ordering:*

(A) *Case (ρ_0, u_0) subsonic:*
  (A.1) *if $\delta > 0$, then $\rho^*_{sup} < \rho_0 < \rho^*_{sub}$;*
  (A.2) *if $\delta \leq 0$, then $\rho^*_{sup} < \rho^*_{sub} \leq \rho_0$.*

(B) *Case (ρ_0, u_0) supersonic:*
  (B.1) *if $\delta \geq 0$, then $\rho^*_{sup} \leq \rho_0 < \rho^*_{sub}$;*
  (B.2) *if $\delta < 0$, then $\rho_0 \leq \rho^*_{sup} < \rho^*_{sub}$.*

(C) *Case (ρ_0, u_0) sonic:*
  $\rho^*_{sup} \leq \rho_0 < \rho^*_{sub}$.

The proof is left to the reader. Now, in order to define a scheme that preserves the nonnegativity of the density, the reconstructed states need to verify $\rho^*_l \leq \rho_l$ and $\rho^*_r \leq \rho_r$ (see Theorem 3.12(i)). We notice that (3.3), (3.4) correspond to the problem (3.18) with, successively, $\delta = Z_l - Z^*$, $\delta = Z_r - Z^*$. Since it is natural to try to choose a solution (ρ^*, u^*) of the same sonicity as (ρ_0, u_0), one sees with the previous lemma that in order to have $\rho^* \leq \rho_0$ one needs $\delta \leq 0$ in the subsonic case, and $\delta \geq 0$ in the supersonic case. Recalling the values $\delta = Z_l - Z^*$ and $\delta = Z_r - Z^*$, this gives that one should have $Z^* \geq \max(Z_l, Z_r)$ in subsonic regions, while $Z^* \leq \min(Z_l, Z_r)$ in supersonic regions. We then observe that it is not possible to satisfy both conditions with Z^* depending continuously on the data. Thus we make the choice of solving exactly the subsonic steady states and disregard supersonic steady states, which justifies $Z^* = \max(Z_l, Z_r)$, i.e., (3.5).

Consider now the function f given by (3.21). For $q \in \mathbb{R}$ fixed, $f(q, \cdot)$ is strictly increasing in $[\rho_s(q), \infty)$ (recall that $\rho_s(q)$ and $m_s(q)$ are defined as (3.14), (3.15)):

$$\begin{aligned} f(q, \cdot)|_{[\rho_s(q),\infty)} : [\rho_s(q), \infty) &\to [m_s(q), \infty), \\ \rho &\mapsto \frac{q^2}{2\rho^2} + \left(e + \frac{p}{\rho}\right)(\rho). \end{aligned} \tag{3.26}$$

We consider its inverse and denote it by $f_r^{-1}(q, \cdot)$:

$$f_r^{-1}(q, \cdot) : [m_s(q), \infty) \to [\rho_s(q), \infty). \tag{3.27}$$

This inverse function corresponds to choosing the subsonic solution to (3.18). With the choice (3.5), equations (3.3), (3.4) correspond to the problem (3.18) with, successively, $\delta = Z_l - Z^* = -(\Delta Z)_+$, $\delta = Z_r - Z^* = -(-\Delta Z)_+$. Therefore, we define the reconstructed states U_l^*, U_r^* by

$$
\begin{aligned}
\rho_l^* &= \min\Big\{\rho_l, f_r^{-1}\Big(\rho_l u_l, \max\Big\{f(\rho_l u_l, \rho_l) - (\Delta Z)_+, m_s(\rho_l u_l)\Big\}\Big)\Big\},\\
u_l^* &= \rho_l u_l/\rho_l^* \quad (u_l^* = u_l \text{ if } \rho_l^* = 0),\\
\rho_r^* &= \min\Big\{\rho_r, f_r^{-1}\Big(\rho_r u_r, \max\Big\{f(\rho_r u_r, \rho_r) - (-\Delta Z)_+, m_s(\rho_r u_r)\Big\}\Big)\Big\},\\
u_r^* &= \rho_r u_r/\rho_r^* \quad (u_r^* = u_r \text{ if } \rho_r^* = 0).
\end{aligned}
\tag{3.28}
$$

Note that these definitions imply that

$$\rho_l^* u_l^* = \rho_l u_l, \qquad \rho_r^* u_r^* = \rho_r u_r. \tag{3.29}$$

LEMMA 3.8. *The definitions* (3.28) *can be interpreted as follows:*
(A) *Case* $\Delta Z \leq 0$*:*
We have the trivial solution to (3.3) $U_l^* = U_l$.
(B) *The general case:*
By Proposition 3.5*, we know that the system* (3.3) *has a solution if and only if*

$$\frac{u_l^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_l) - (\Delta Z)_+ \geq m_s(\rho_l u_l). \tag{3.30}$$

(B.1) *If* (ρ_l, u_l) *is a supersonic point or a sonic point, then* $U_l^* = U_l$.
(B.2) *If* (ρ_l, u_l) *is a subsonic point and we have strict inequality in* (3.30)*, then* (ρ_l^*, u_l^*) *is the subsonic solution to* (3.3).
(B.3) *If* (ρ_l, u_l) *is a subsonic point and we have equality in* (3.30) *or the inequality is not satisfied, then* $\rho_l^* = \rho_s(\rho_l u_l)$.
Similar statements hold for (ρ_r^*, u_r^*).

The case $\rho_l^* = 0$ could pose some problems in the previous definition of u_l^*, but as we consider conservative variables, the product $\rho_l^* u_l^*$ is well defined. The following result shows that there is indeed no problem of continuity.

LEMMA 3.9. *The reconstructed states* (3.28) *verify the following:*
(i)

$$\min\Big\{\rho_l, \rho_s(\rho_l u_l)\Big\} \leq \rho_l^* \leq \rho_l, \qquad \min\Big\{\rho_r, \rho_s(\rho_r u_r)\Big\} \leq \rho_r^* \leq \rho_r; \tag{3.31}$$

(ii) *independently of the other arguments, one has*

$$\lim_{\rho_l \to 0} \rho_l^* = 0, \qquad \lim_{\rho_r \to 0} \rho_r^* = 0; \tag{3.32}$$

(iii) *for* $\rho_l \geq 0$*,* $\rho_r \geq 0$*,* ΔZ *fixed, we have*

$$\lim_{u_l \to 0} \rho_l^* = \left(e + \frac{p}{\rho}\right)^{-1}\left(\max\left\{\left(e + \frac{p}{\rho}\right)(\rho_l) - (\Delta Z)_+, m_s(0)\right\}\right), \tag{3.33}$$

$$\lim_{u_r \to 0} \rho_r^* = \left(e + \frac{p}{\rho}\right)^{-1}\left(\max\left\{\left(e + \frac{p}{\rho}\right)(\rho_r) - (-\Delta Z)_+, m_s(0)\right\}\right); \tag{3.34}$$

(iv) *for ρ_l, ρ_r bounded, one has*

$$\lim_{u_l \to 0} u_l^* = 0, \qquad \lim_{u_r \to 0} u_r^* = 0. \tag{3.35}$$

Proof. Note that (iii) means the continuity of ρ_l^* and ρ_r^* in this asymptotics. We shall give only the proof for ρ_l^*, the proof for ρ_r^* being similar.

According to Lemma 3.8, the only case when (3.31) is nontrivial is (B.2) with $\Delta Z \geq 0$. Then Corollary 3.6 allows one to conclude the proof of (i). Then, (ii) is a consequence of (i).

In order to prove (iii), for $\rho_l \geq 0$ and ΔZ fixed, we shall denote

$$\begin{aligned} \beta_l(u_l) &= \max\Big\{f(\rho_l u_l, \rho_l) - (\Delta Z)_+, m_s(\rho_l u_l)\Big\}, \\ \alpha_l(u_l) &= f_r^{-1}(\rho_l u_l, \beta_l(u_l)). \end{aligned} \tag{3.36}$$

We have

$$\lim_{u_l \to 0} \beta_l(u_l) = \max\left\{\left(e + \frac{p}{\rho}\right)(\rho_l) - (\Delta Z)_+, m_s(0)\right\}. \tag{3.37}$$

Since $\beta_l(u_l) \geq m_s(\rho_l u_l)$, we have $\alpha_l(u_l) \geq \rho_s(\rho_l u_l)$, which yields

$$0 \leq \frac{(\rho_l u_l)^2}{\alpha_l(u_l)^2} \leq \frac{(\rho_l u_l)^2}{\rho_s(\rho_l u_l)^2} = p'\Big(\rho_s(\rho_l u_l)\Big). \tag{3.38}$$

Therefore, $(\rho_l u_l)^2/\alpha_l(u_l)^2$ tends to 0 as $u_l \to 0$. Then, using the identity

$$\frac{(\rho_l u_l)^2}{2(\alpha_l(u_l))^2} + \left(e + \frac{p}{\rho}\right)(\alpha_l(u_l)) = \beta_l(u_l), \tag{3.39}$$

we get

$$\lim_{u_l \to 0} \alpha_l(u_l) = \left(e + \frac{p}{\rho}\right)^{-1}\left(\max\left\{\left(e + \frac{p}{\rho}\right)(\rho_l) - (\Delta Z)_+, m_s(0)\right\}\right). \tag{3.40}$$

Since the right-hand side is at most ρ_l, we deduce that $\rho_l^* = \min\{\rho_l, \alpha_l(u_l)\}$ has the same limit, which concludes the proof of (iii).

The last statement (iv) is also a consequence of (i) since either $\rho_l^* \geq \rho_l$, giving $|u_l^*| \leq |u_l|$, or $\rho_l^* \geq \rho_s(\rho_l u_l)$, giving

$$(u_l^*)^2 = \frac{(\rho_l u_l)^2}{(\rho_l^*)^2} \leq \frac{(\rho_l u_l)^2}{\rho_s(\rho_l u_l)^2} = p'\Big(\rho_s(\rho_l u_l)\Big). \tag{3.41}$$

Since the right-hand side tends to 0, this concludes the proof. □

We would like to end this subsection by giving an iterative procedure in order to solve the system (3.18). According to Lemma 3.8, we need to solve this system only in the case (B.2) with $\Delta Z > 0$. This is done as follows.

PROPOSITION 3.10. *Let $u_0 \in \mathbb{R}\backslash\{0\}$, $\rho_0 > 0$, and $\delta < 0$. We suppose that $u_0^2 < p'(\rho_0)$ and that (3.19) is strictly satisfied. Let $V_0 = \frac{u_0^2}{2} + (e + \frac{p}{\rho})(\rho_0) + \delta$, and let $\psi(\rho) = \rho^\alpha(f(q_0, \rho) - V_0)$, where f is the function given by (3.21), and $q_0 = \rho_0 u_0$. Then, for $\alpha \geq 3/2$, the relation*

$$\rho_{n+1} = \rho_n - \frac{\psi(\rho_n)}{\psi'(\rho_n)} \tag{3.42}$$

(starting from ρ_0) defines a decreasing sequence that converges to ρ^*_{sub}, the subsonic solution to (3.18).

Proof. According to Proposition 3.5, there is a unique solution $\rho^*_{sub} \in (\rho_s, \rho_0)$ to the equation $\psi(\rho) = 0$. We have

$$
\psi'(\rho) = \alpha\rho^{\alpha-1}\Big(f(q_0,\rho) - V_0\Big) + \rho^\alpha \frac{\partial f}{\partial \rho}(q_0,\rho),
$$

$$
\psi''(\rho) = \alpha(\alpha-1)\rho^{\alpha-2}\Big(f(q_0,\rho) - V_0\Big) + 2\alpha\rho^{\alpha-1}\frac{\partial f}{\partial \rho}(q_0,\rho) + \rho^\alpha \frac{\partial^2 f}{\partial \rho^2}(q_0,\rho), \tag{3.43}
$$

and, for $\rho \geq \rho^*_{sub}$,

$$
\begin{aligned}
& f(q_0,\rho) - V_0 \geq 0,\\
& \frac{\partial f}{\partial \rho}(q_0,\rho) = \frac{\rho^2 p'(\rho) - q_0^2}{\rho^3} > 0,\\
& \frac{\partial^2 f}{\partial \rho^2}(q_0,\rho) = -\frac{3}{\rho}\frac{\partial f}{\partial \rho}(q_0,\rho) + \frac{1}{\rho^3}\frac{\partial}{\partial \rho}\Big(\rho^2 p'(\rho)\Big) \geq -\frac{3}{\rho}\frac{\partial f}{\partial \rho}(q_0,\rho).
\end{aligned} \tag{3.44}
$$

Thus, for $\alpha \geq 3/2$, ψ is strictly increasing and convex. Therefore, the Newton method converges to the zero of the function, and this proves the result. □

3.3. Definition of left and right fluxes. In order to define the left and right numerical fluxes as (3.1), taking the definitions (3.28) for U_l^*, U_r^* (recall that Z^* is given by (3.5)), we still need to define T_l and T_r. We denote $(\mathcal{F}^\rho, \mathcal{F}^{\rho u}) \equiv \mathcal{F}(U_l^*, U_r^*)$ and define

$$
\begin{aligned}
T_l(U_l,U_r,\Delta Z) ={}& \frac{\rho_l - \rho_l^*}{\rho_l^*}\Big(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^* \mathcal{F}^\rho\Big) - (u_l^* - u_l)\,\mathcal{F}^\rho\\
&+ \left(\left(e + \frac{p}{\rho}\right)(\rho_l^*) - \left(e + \frac{p}{\rho}\right)(\rho_l) + (\Delta Z)_+ + \frac{(u_l^*)^2}{2} - \frac{u_l^2}{2}\right)\frac{\mathcal{F}^\rho}{u_l},
\end{aligned} \tag{3.45}
$$

$$
\begin{aligned}
T_r(U_l,U_r,\Delta Z) ={}& \frac{\rho_r - \rho_r^*}{\rho_r^*}\Big(\mathcal{F}^{\rho u} - p(\rho_r^*) - u_r^* \mathcal{F}^\rho\Big) - (u_r^* - u_r)\,\mathcal{F}^\rho\\
&+ \left(\left(e + \frac{p}{\rho}\right)(\rho_r^*) - \left(e + \frac{p}{\rho}\right)(\rho_r) + (-\Delta Z)_+ + \frac{(u_r^*)^2}{2} - \frac{u_r^2}{2}\right)\frac{\mathcal{F}^\rho}{u_r}.
\end{aligned} \tag{3.46}
$$

The definition (3.45) (resp., (3.46)) is ambiguous in the case when $u_l = 0$ or $\rho_l^* = 0$ (resp., when $u_r = 0$ or $\rho_r^* = 0$). In order to overcome this difficulty, here we make the convention that "$0/0 = 0$." Then, one has to take into account the following remarks. They are stated for T_l, but of course similar statements hold for T_r.

1. If (ρ_l^*, u_l^*) solves the system (3.3) (with (3.5)), the factor of $\frac{\mathcal{F}^\rho}{u_l}$ in T_l vanishes.

2. If (ρ_l, u_l) is a supersonic point, we have $U_l^* = U_l$ and $T_l \equiv \frac{\mathcal{F}^\rho}{u_l}(\Delta Z)_+$, which is well defined since $u_l^2 > p'(\rho_l) \geq 0$.

3. If (ρ_l, u_l) is a subsonic point and $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - (\Delta Z)_+ > m_s(\rho_l u_l)$, according to (B.2) in Lemma 3.8 we have that $\rho_l^* > 0$, and the factor of $\frac{\mathcal{F}^\rho}{u_l}$ in T_l vanishes; thus T_l is well defined. This is true in particular when (ρ_l, u_l) is a subsonic point if we

consider a continuous bottom $z(x)$, which implies that ΔZ is small for a sufficiently fine grid.

4. If (ρ_l, u_l) is a subsonic point and $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - (\Delta Z)_+ \le m_s(\rho_l u_l)$, but $u_l \neq 0$, according to (B.3) in Lemma 3.8 we have $\rho_l^* = \rho_s(\rho_l u_l) > 0$; thus T_l is well defined.

5. Some difficulties may arise for (ρ_l, u_l) sonic close to $(0,0)$, or for (ρ_l, u_l) subsonic with $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - (\Delta Z)_+ \le m_s(\rho_l u_l)$ and u_l close to 0 (which also implies that (ρ_l, u_l) is close to $(0,0)$). In these cases, the flux $\mathcal{F}$ has to verify some conditions in order to define $\frac{\mathcal{F}^\rho}{u_l}$ and $(\rho_l^*)^{-1}\big(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^* \mathcal{F}^\rho\big)$. As $\mathcal{F}$ is consistent with F, we expect these quantities to be unambiguously defined.

LEMMA 3.11. *The definitions* (3.45), (3.46) *of* T_l *and* T_r *imply that the conditions* (3.12) *are satisfied (with* Z^* *given by* (3.5)*).*

Proof. Consider the case of W_l. We have

$$u_l \frac{\rho_l - \rho_l^*}{\rho_l^*} = u_l^* - u_l; \tag{3.47}$$

thus

$$\begin{aligned} u_l T_l &= \Big(u_l^* - u_l\Big)\Big(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^* \mathcal{F}^\rho - u_l \mathcal{F}^\rho\Big) \\ &\quad + \left(\left(e + \frac{p}{\rho}\right)(\rho_l^*) - \left(e + \frac{p}{\rho}\right)(\rho_l) + (\Delta Z)_+ + \frac{(u_l^*)^2}{2} - \frac{u_l^2}{2}\right)\mathcal{F}^\rho \\ &= \Big(u_l^* - u_l\Big)\Big(\mathcal{F}^{\rho u} - p(\rho_l^*)\Big) \\ &\quad + \left(\left(e + \frac{p}{\rho}\right)(\rho_l^*) - \left(e + \frac{p}{\rho}\right)(\rho_l) + (\Delta Z)_+ - \frac{(u_l^*)^2}{2} + \frac{u_l^2}{2}\right)\mathcal{F}^\rho. \end{aligned} \tag{3.48}$$

Putting this value in (3.10) gives $W_l = 0$. ☐

3.4. Properties of the subsonic reconstruction scheme.

THEOREM 3.12. *Let* $\mathcal{F}(U_l, U_r)$ *be a given consistent numerical flux for the Saint-Venant problem without source that preserves the nonnegativity of* ρ *by interface and satisfies a semidiscrete entropy inequality for the entropy pair* (η, G) *given by* (2.8)*. Then the scheme* (2.12), (2.13), *with numerical fluxes* (3.1), (3.28), (3.45), (3.46),

(0) *is conservative in density,*

(i) *preserves the nonnegativity of* ρ *by interface,*

(ii) *preserves the discrete subsonic steady states,*

(iii) *is consistent with the Saint-Venant system away from sonic points, and*

(iv) *satisfies a semidiscrete entropy inequality associated with the entropy pair* $(\widetilde{\eta}, \widetilde{G})$ *in* (2.8)*.*

Proof. Notice first that when $\Delta Z = 0$ we have $U_l^* = U_l$, $U_r^* = U_r$, $T_l = 0$, and $T_r = 0$, so that the scheme (3.1) reduces to the conservative scheme with numerical flux $\mathcal{F}$.

Property (0) is obvious from the definition (3.1) and the characterization (2.15).

For (i), the assumption that $\mathcal{F}(U_l, U_r)$ preserves nonnegativity of ρ by interface means that (see [7]) there exists some $\sigma_l(U_l, U_r) < 0 < \sigma_r(U_l, U_r)$ such that

$$\rho_l + \frac{\mathcal{F}^\rho(U_l, U_r) - \rho_l u_l}{\sigma_l(U_l, U_r)} \ge 0, \qquad \rho_r + \frac{\mathcal{F}^\rho(U_l, U_r) - \rho_r u_r}{\sigma_r(U_l, U_r)} \ge 0 \tag{3.49}$$

for any U_l and U_r (with nonnegative densities ρ_l, ρ_r). This implies in particular that

$$\rho_l^* + \frac{\mathcal{F}^\rho(U_l^*, U_r^*) - \rho_l^* u_l^*}{\sigma_l(U_l^*, U_r^*)} \geq 0, \qquad \rho_r^* + \frac{\mathcal{F}^\rho(U_l^*, U_r^*) - \rho_r^* u_r^*}{\sigma_r(U_l^*, U_r^*)} \geq 0. \tag{3.50}$$

According to (3.28), one has $\rho_l^* u_l^* = \rho_l u_l$, $\rho_r^* u_r^* = \rho_r u_r$, and $\rho_l^* \leq \rho_l$, $\rho_r^* \leq \rho_r$; thus

$$\rho_l + \frac{\mathcal{F}^\rho(U_l^*, U_r^*) - \rho_l u_l}{\sigma_l(U_l^*, U_r^*)} \geq 0, \qquad \rho_r + \frac{\mathcal{F}^\rho(U_l^*, U_r^*) - \rho_r u_r}{\sigma_r(U_l^*, U_r^*)} \geq 0, \tag{3.51}$$

proving that the scheme preserves the nonnegativity of ρ by interface. The associated speeds involved in the CFL condition are $\sigma_l(U_l^*, U_r^*)$, $\sigma_r(U_l^*, U_r^*)$.

In order to prove (ii), consider left and right states U_l, U_r such that the steady state equations (2.10) are satisfied,

$$\begin{cases} \dfrac{u_l^2}{2} + \left(e + \dfrac{p}{\rho}\right)(\rho_l) + Z_l = \dfrac{u_r^2}{2} + \left(e + \dfrac{p}{\rho}\right)(\rho_r) + Z_r, \\ \rho_l u_l = \rho_r u_r, \end{cases} \tag{3.52}$$

and such that both are subsonic:

$$u_l^2 < p'(\rho_l), \qquad u_r^2 < p'(\rho_r). \tag{3.53}$$

Note in particular that $\rho_l > 0$, $\rho_r > 0$. Recall that $\Delta Z \equiv Z_r - Z_l$. If $U_l = U_r$, then $\Delta Z = 0$ and according to the remark above one has $\mathcal{F}_l = \mathcal{F}(U_l, U_r) = F(U_l)$, $\mathcal{F}_r = \mathcal{F}(U_l, U_r) = F(U_r)$, proving (2.18). Assume now that $U_l \neq U_r$. Then $\Delta Z \neq 0$; otherwise (3.52) would give two subsonic solutions to a system (3.18), which is not possible by Proposition 3.5. Consider first the case when $\Delta Z > 0$. Then according to Lemma 3.8 we have

$$U_l^* = U_r^* = U_r \tag{3.54}$$

and

$$\mathcal{F}(U_l^*, U_r^*) = F(U_r), \qquad \mathcal{F}^{\rho u}(U_l^*, U_r^*) - p(\rho_l^*) - u_l^* \mathcal{F}^\rho(U_l^*, U_r^*) = 0; \tag{3.55}$$

thus

$$T_l = -(u_l^* - u_l)\rho_l^* u_l^*, \qquad T_r = 0. \tag{3.56}$$

According to the remark after (3.2), the relations (2.18) are satisfied. The case $\Delta Z < 0$ is similar, with $U_l^* = U_r^* = U_l$. This proves (ii).

Property (iv) follows from Lemmas 3.1, 3.2, and 3.11.

It remains to prove the consistency (iii). The first property (2.16) is obvious according to the remark above on the case $\Delta Z = 0$. Regarding (2.17), taking into account (3.1), we have to prove that

$$p(\rho_r) - p(\rho_r^*) + T_r - p(\rho_l) + p(\rho_l^*) - T_l = -\rho \Delta Z + o(\Delta Z) \tag{3.57}$$

as $U_l, U_r \to U$ and $\Delta Z \to 0$. As stated, we consider only the case when U is not sonic. Let us assume that $\Delta Z \geq 0$, the complementary case being similar. Then $U_r^* = U_r$ and $T_r = 0$.

(a) Case (ρ, u) supersonic. Then (ρ_l, u_l) is also supersonic if it is close enough to (ρ, u) (and in particular $u_l \neq 0$), and we have

$$\rho_l^* = \rho_l, \quad u_l^* = u_l, \quad \rho_r^* = \rho_r, \quad u_r^* = u_r, \tag{3.58}$$

$$\mathcal{F}_r^{\rho u} - \mathcal{F}_l^{\rho u} = -T_l = -\mathcal{F}^\rho(U_l, U_r)\frac{\Delta Z}{u_l} = -\rho\Delta Z + o(\Delta Z). \tag{3.59}$$

(b) Case (ρ, u) subsonic. Then $\rho > 0$, and $\frac{u^2}{2} + (e + \frac{p}{\rho})(\rho) > m_s(\rho u)$. Therefore, for U_l close enough to U and ΔZ small enough, we have (ρ_l, u_l) subsonic, $\rho_l > 0$, $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - \Delta Z > m_s(\rho_l u_l)$. From (B.2) in Lemma 3.8 we have that $\rho_l^* > 0$, and we compute

$$\begin{aligned} \rho_l^* &= f_r^{-1}\left(\rho_l u_l, \frac{u_l^2}{2} + \left(e + \frac{p}{\rho}\right)(\rho_l) - \Delta Z\right) \\ &= f_r^{-1}\Big(\rho_l u_l, f(\rho_l u_l, \rho_l) - \Delta Z\Big) \\ &= \rho_l + O(\Delta Z), \end{aligned} \tag{3.60}$$

$$u_l^* = \frac{\rho_l u_l}{\rho_l^*} = u_l + O(\Delta Z). \tag{3.61}$$

Now, let us denote $\phi(\rho) = (e + \frac{p}{\rho})(\rho)$. Since ϕ is a strictly increasing function, we can consider its inverse ϕ^{-1}. According to (3.60),

$$\left(e + \frac{p}{\rho}\right)(\rho_l^*) = \left(e + \frac{p}{\rho}\right)(\rho_l) + \frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} - \Delta Z; \tag{3.62}$$

thus using (3.61) we get

$$\begin{aligned} \rho_l^* &= \phi^{-1}\left(\phi(\rho_l) + \frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} - \Delta Z\right) \\ &= \rho_l + (\phi^{-1})'(\phi(\rho_l))\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} - \Delta Z\right) + O\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} - \Delta Z\right)^2 \\ &= \rho_l + \frac{\rho_l}{p'(\rho_l)}\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} - \Delta Z\right) + O(\Delta Z)^2. \end{aligned} \tag{3.63}$$

Then,

$$\begin{aligned} p(\rho_l^*) &= p(\rho_l) + p'(\rho_l)(\rho_l^* - \rho_l) + O(\Delta Z)^2 \\ &= p(\rho_l) - \rho_l\Delta Z + \rho_l\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2}\right) + O(\Delta Z)^2. \end{aligned} \tag{3.64}$$

We also have

$$\frac{\rho_l - \rho_l^*}{\rho_l^*}\Big(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^*\mathcal{F}^\rho\Big) = O(\Delta Z)\cdot o(1) = o(\Delta Z). \tag{3.65}$$

Therefore,

$$
\begin{aligned}
\mathcal{F}_r^{\rho u} - \mathcal{F}_l^{\rho u} &= p(\rho_l^*) - p(\rho_l) - T_l \\
&= p(\rho_l^*) - p(\rho_l) - \frac{\rho_l - \rho_l^*}{\rho_l^*}\left(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^* \mathcal{F}^\rho\right) + (u_l^* - u_l)\,\mathcal{F}^\rho \\
&= -\rho\Delta Z + \rho_l\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2}\right) + (u_l^* - u_l)\,\mathcal{F}^\rho + o(\Delta Z) \qquad (3.66)\\
&= -\rho\Delta Z + (u_l - u_l^*)\left(\rho_l \frac{u_l + u_l^*}{2} - \mathcal{F}^\rho\right) + o(\Delta Z) \\
&= -\rho\Delta Z + o(\Delta Z),
\end{aligned}
$$

which yields (iii). ∎

3.5. Comments on the consistency at sonic points. The proof of (iii) in the previous theorem when we are close to a sonic point (ρ, u) involves some problems. Consider the case when $U_l,\ U_r \to U,\ \Delta Z \to 0$, with $u^2 = p'(\rho) \neq 0$. Assume as previously that $\Delta Z \geq 0$. If (ρ_l, u_l) is supersonic, then (3.59) is valid.

Otherwise, if (ρ_l, u_l) is subsonic and $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - \Delta Z \geq m_s(\rho_l u_l)$, the computations made in the proof of consistency in the subsonic case can be followed, with the result

$$
\mathcal{F}_r^{\rho u} - \mathcal{F}_l^{\rho u} = -\rho\Delta Z + O(\Delta Z) + O\Big(|\Delta Z|^{1/2}|U_l - U_r|\Big). \qquad (3.67)
$$

Assume now that either (ρ_l, u_l) is sonic, or (ρ_l, u_l) is subsonic with $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - \Delta Z < m_s(\rho_l u_l)$. In any case we have $\frac{u_l^2}{2} + (e + \frac{p}{\rho})(\rho_l) - \Delta Z \leq m_s(\rho_l u_l)$; thus $\rho_l^* = \rho_s(\rho_l u_l) > 0$, and since

$$
0 \leq f(\rho_l u_l, \rho_l) - m_s(\rho_l u_l) \leq \Delta Z \qquad (3.68)
$$

and $\rho_s(\rho_l u_l)$ is the minimum point of the function $f(\rho_l u_l, \cdot)$, we deduce that

$$
\rho_s(\rho_l u_l) - \rho_l = O(|\Delta Z|^{1/2}), \qquad u_l^* = \frac{\rho_l u_l}{\rho_s(\rho_l u_l)} = u_l + O(|\Delta Z|^{1/2}). \qquad (3.69)
$$

Thus, assuming that the numerical flux $\mathcal{F}$ is Lipschitz continuous,

$$
\frac{\rho_l - \rho_l^*}{\rho_l^*}\left(\mathcal{F}^{\rho u} - p(\rho_l^*) - u_l^* \mathcal{F}^\rho\right) = O(\Delta Z) + O\Big(|\Delta Z|^{1/2}|U_l - U_r|\Big). \qquad (3.70)
$$

Now, from (3.68), we deduce that $(e + \frac{p}{\rho})(\rho_l^*) = (e + \frac{p}{\rho})(\rho_l) + \frac{u_l^2}{2} - \frac{(u_l^*)^2}{2} + O(\Delta Z)$, and using calculations analogous to the ones applied in the proof of the consistency in the subsonic case, one gets

$$
\begin{aligned}
\rho_l^* &= \rho_l + \frac{\rho_l}{p'(\rho_l)}\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2}\right) + O(\Delta Z), \\
p(\rho_l^*) &= p(\rho_l) + \rho_l\left(\frac{u_l^2}{2} - \frac{(u_l^*)^2}{2}\right) + O(\Delta Z)
\end{aligned} \qquad (3.71)
$$

and

$$
\begin{aligned}
\mathcal{F}_r^{\rho u} - \mathcal{F}_l^{\rho u} &= p(\rho_l^*) - p(\rho_l) - T_l \\
&= p(\rho_l^*) - p(\rho_l) + (u_l^* - u_l)\mathcal{F}^\rho + O(\Delta Z) + O\Big(|\Delta Z|^{1/2}|U_l - U_r|\Big) \\
&= (u_l - u_l^*)\Big(\rho_l \frac{u_l + u_l^*}{2} - \mathcal{F}^\rho\Big) + O(\Delta Z) + O\Big(|\Delta Z|^{1/2}|U_l - U_r|\Big) \\
&= O(\Delta Z) + O\Big(|\Delta Z|^{1/2}|U_l - U_r|\Big).
\end{aligned}
\tag{3.72}
$$

Thus, in any case, (3.67) is valid.

This property (3.67) does not mean consistency in the sense of (2.17), but anyway, one can make the following observation. Assume that, at the point considered, one has $dZ/dx = 0$. Then $\Delta Z = o(\Delta x)$, $U_r - U_l = O(\Delta x)$, and therefore (3.67) yields

$$
\mathcal{F}_r^{\rho u} - \mathcal{F}_l^{\rho u} = o(\Delta x),
\tag{3.73}
$$

which means consistency with the vanishing source. Since the condition $dZ/dx = 0$ is generically satisfied at sonic points (see [10]), it justifies the global consistency of the scheme, except maybe close to the point $(\rho, u) = (0, 0)$. We shall see in the numerical computations that even if the numerical fluxes can sometimes take large values, the scheme behaves reasonably well in the presence of data close to $(\rho, u) = (0, 0)$.

4. Extension to the two-dimensional case. The proposed scheme can be easily adapted to the two-dimensional (2D) case by using the usual approach of a finite volume method for multidimensional systems. We describe here the main ideas and refer the reader to [7] and the references therein for further details.

Consider the system

$$
\partial_t U + \partial_1 F_1(U, Z) + \partial_2 F_2(U, Z) + B_1(U, Z)\partial_1 Z + B_2(U, Z)\partial_2 Z = 0.
\tag{4.1}
$$

We discretize the space domain by means of a mesh made of cells C_i, the control volumes. The interface Γ_{ij} between two cells C_i and C_j is assumed to be a segment, and we denote by n_{ij} the unit normal vector oriented from C_i to C_j. Then, define the numerical scheme

$$
U_i^{n+1} = U_i^n - \frac{\Delta t}{|C_i|} \sum_{j \in K_i} |\Gamma_{ij}| F_{ij} = 0,
\tag{4.2}
$$

where U_i represents the cell-average of U over C_i, $|C_i|$ is the volume of C_i, $|\Gamma_{ij}|$ is the length of Γ_{ij}, and $F_{ij} = F(U_i, U_j, Z_i, Z_j, n_{ij})$ is a numerical flux approximating $n^1 F_1 + n^2 F_2$ which gives the exchange term between C_j and C_i.

Now, consider the particular case of 2D shallow water equations

$$
\begin{cases}
\partial_t \rho + \partial_1(\rho u) + \partial_2(\rho v) = 0, \\
\partial_t(\rho u) + \partial_1(\rho u^2 + p(\rho)) + \partial_2(\rho u v) + \rho g \partial_1 z = 0, \\
\partial_t(\rho v) + \partial_1(\rho u v) + \partial_2(\rho v^2 + p(\rho)) + \rho g \partial_2 z = 0,
\end{cases}
\tag{4.3}
$$

where $\rho(t, x) \geq 0$ is the water height, $u(t, x), v(t, x)$ are the two components of the velocity field, $p(\rho)$ is the pressure, and $z(x)$ is the topography.

This system has the property of being invariant by rotation, which enables us to easily define a numerical flux for the 2D system from a given one-dimensional (1D) flux (see [7]).

Thus, the fluxes at each interface can be computed with the numerical fluxes defined in the previous section after an appropriate rotation.

The scheme will then automatically preserve the nonnegativity of the water height and satisfy a semidiscrete entropy inequality.

But concerning well-balancing, it follows from (4.1) that steady states that can be preserved by such a method are those that can be seen as interface by interface steady states. In particular, the 2D scheme will be able to preserve any lake-at-rest steady state but not all subsonic equilibria. Only those equilibria that correspond to the rotation of a 1D subsonic steady state shall be preserved (provided that we choose a rectangular mesh in the direction of the fluid). But any more general steady state where the dependence in x_1 and x_2 is nontrivially balanced cannot be preserved.

5. Application to the Euler–Poisson system. Let us consider the Euler–Poisson system

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) = -\rho \partial_x \phi, \\ -\partial^2_{xx}\phi = \rho - \rho_b, \end{cases} \tag{5.1}$$

where $\rho_b \equiv \rho_b(x) \geq 0$ is given. The system (5.1) is set for $t > 0$, $0 < x < l$, with initial and boundary conditions

$$\begin{cases} \rho(t=0,\cdot) = \rho_b, \\ u(t=0,\cdot) = u_0, \\ \phi(l) - \phi(0) = V, \\ \rho u(t, x=0) = q_0 \geq 0, \\ \rho(t, x=l) = \rho_b(l). \end{cases} \tag{5.2}$$

As usual, one has to complete (5.1) by an entropy inequality. In order to describe the steady states, we subtract u times the first equation in (5.1) from the second, divide the result by ρ, and get

$$\partial_t u + \partial_x(u^2/2 + e(\rho) + p(\rho)/\rho + \phi) = 0. \tag{5.3}$$

Therefore, the steady states are determined by the relations

$$\begin{cases} \rho u = Cst, \\ \frac{u^2}{2} + e(\rho) + \frac{p(\rho)}{\rho} + \phi = Cst. \end{cases} \tag{5.4}$$

We can observe that the Euler–Poisson system is of the type (2.1), where the bottom Z is replaced by a function ϕ that is time dependent.

The subsonic reconstruction scheme can be applied to this system by "freezing" the potential on a time interval as follows. Given an approximation of ϕ at time t_n, $\phi^n = \phi(t_n,\cdot)$, we solve the system

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) = -\rho \partial_x \phi^n \end{cases} \tag{5.5}$$

in the interval $[t_n, t_{n+1})$ using the subsonic reconstruction scheme, where ϕ^n stands for Z. We obtain approximations ρ^{n+1}, u^{n+1}. Finally, we solve the ODE

$$-\partial^2_{xx}\phi^{n+1} = \rho^{n+1} - \rho_b \tag{5.6}$$

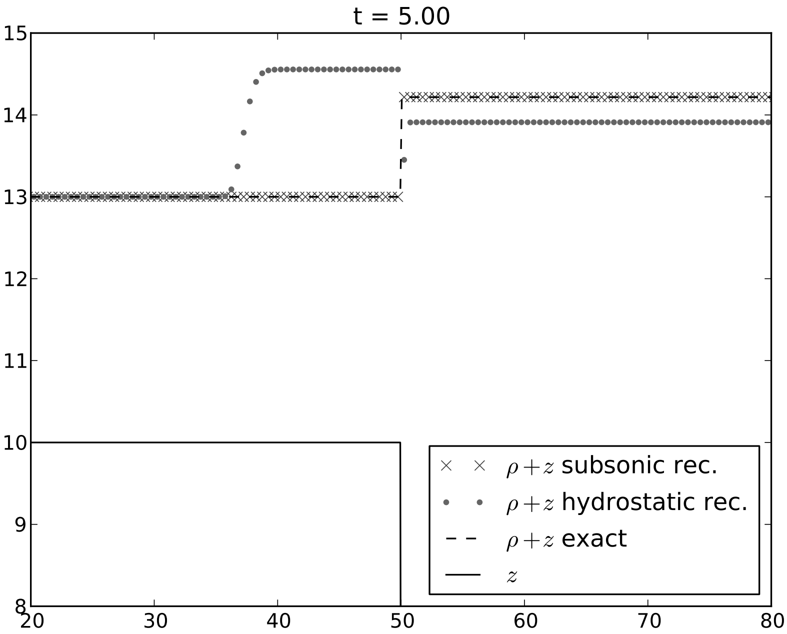

FIG. 2. *Subsonic steady state; $\rho + z$ at $t = 5$.*

in order to get the new potential. It is obvious that this algorithm is well-balanced since the freezing of the potential does not introduce any error in the case of a steady state.

The interest of the subsonic reconstruction scheme in this context is the ability to compute with high accuracy flows which are close to a subsonic steady state with constant discharge $\rho u \neq 0$.

6. Numerical results.

6.1. Saint-Venant system. In order to evaluate our method, we compare the subsonic reconstruction scheme described here to the original hydrostatic reconstruction scheme of [1]. We use first-order resolution, with the CFL 1 condition induced by the nonnegativity of density (see the proof of (i) in Theorem 3.12). A second-order extension can be used as in [1], and in that case no major differences are observed between the two reconstructions. The numerical flux $\mathcal{F}$ chosen here is the one obtained from the Suliciu relaxation system described in [7]. We take $p(\rho) = g\rho^2/2$, $g = 9.81$ and use 200 points in the considered interval in each case.

We consider first a subsonic steady state in the interval $(0, 100)$. The initial data are given by

$$\rho_0(x) = \begin{cases} 3 & \text{if } x \leq 50, \\ 14.2175 & \text{if } x > 50, \end{cases} \qquad u_0(x) = \begin{cases} 5 & \text{if } x \leq 50, \\ 1.055 & \text{if } x > 50, \end{cases} \tag{6.1}$$

$$z(x) = \begin{cases} 10 & \text{if } x \leq 50, \\ 0 & \text{if } x > 50. \end{cases} \tag{6.2}$$

The results are shown in Figures 2 and 3. As we see, the subsonic reconstruction scheme maintains the subsonic steady state, as we have proved. This is an improvement with respect to the hydrostatic reconstruction scheme, which does not.

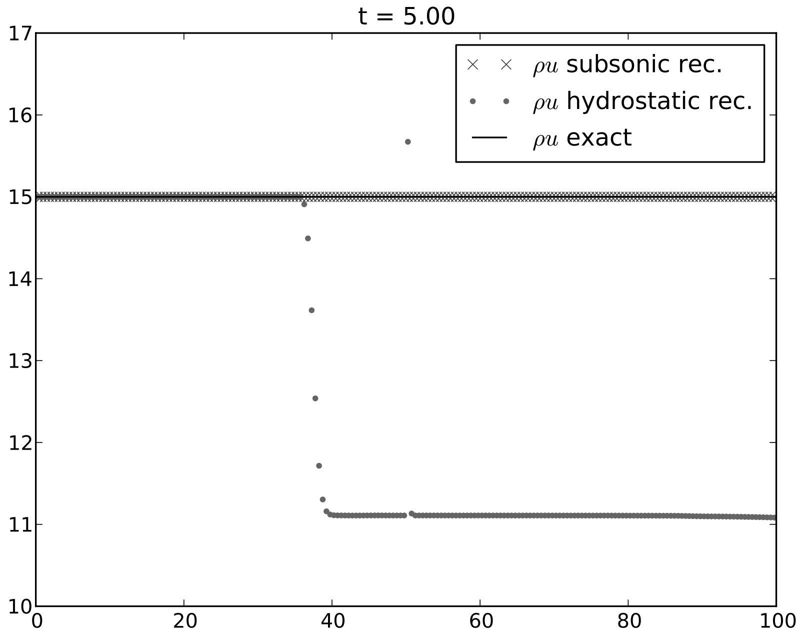

FIG. 3. *Subsonic steady state; ρu at $t = 5$.*

Then, we consider a supersonic steady state, still in the interval $(0, 100)$:

$$\rho_0(x) = \begin{cases} 2 & \text{if } x \le 50, \\ 0.635 & \text{if } x > 50, \end{cases} \qquad u_0(x) = \begin{cases} 5 & \text{if } x \le 50, \\ 15.7474 & \text{if } x > 50, \end{cases} \tag{6.3}$$

$$z(x) = \begin{cases} 10 & \text{if } x \le 50, \\ 0 & \text{if } x > 50. \end{cases} \tag{6.4}$$

We see in Figures 4, 5, 6, and 7 that neither of the two schemes gives the right solution, but the subsonic reconstruction scheme is more accurate.

We consider now a classical transcritical shock test. The space domain is $(0, 25)$, the initial data are $\rho_0(x) = 0.33$, $u_0(x) = 0.18/0.33$, and the topography is

$$z(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8 < x < 12, \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

The boundary conditions are taken as $\rho u(x = 0) = 0.18$ and $\rho(x = 25) = 0.33$. The results are shown in Figures 8 and 9. We see that the subsonic reconstruction scheme gives a solution which is sharper on the left part of the discontinuity than the hydrostatic reconstruction.

6.2. Euler–Poisson system. We solve the Euler–Poisson system (5.1), (5.2) with $p(\rho) = \kappa\rho^\gamma$,

$$\kappa = 1, \quad \gamma = 1.1, \tag{6.6}$$

for $x \in (0, 0.6)$ and with initial and boundary conditions

$$\begin{gathered} \rho_b(x) = \begin{cases} 1 & \text{if } x \in (0.1, 0.5), \\ 100 & \text{otherwise,} \end{cases} \\ u_0 = 0, \quad V = -1. \end{gathered} \tag{6.7}$$

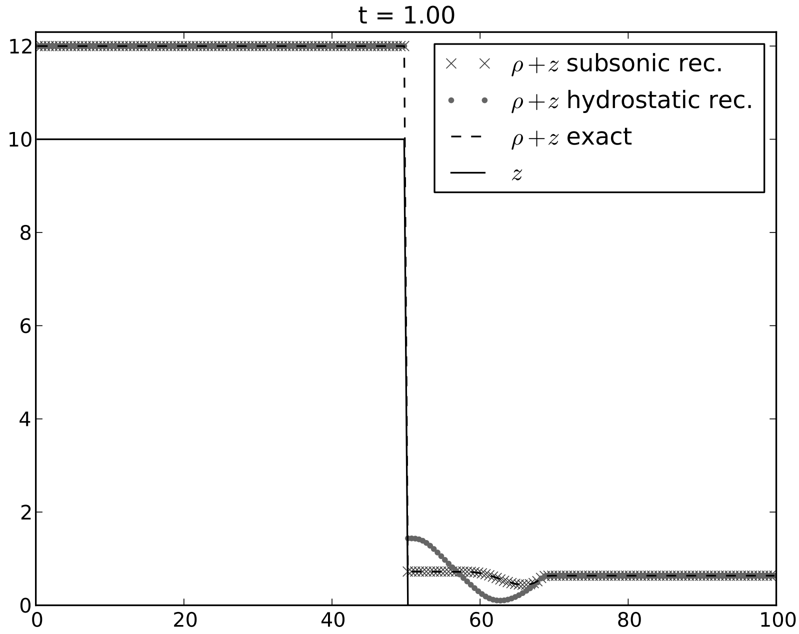

Fig. 4. *Supersonic steady state; $\rho + z$ at $t = 1$.*

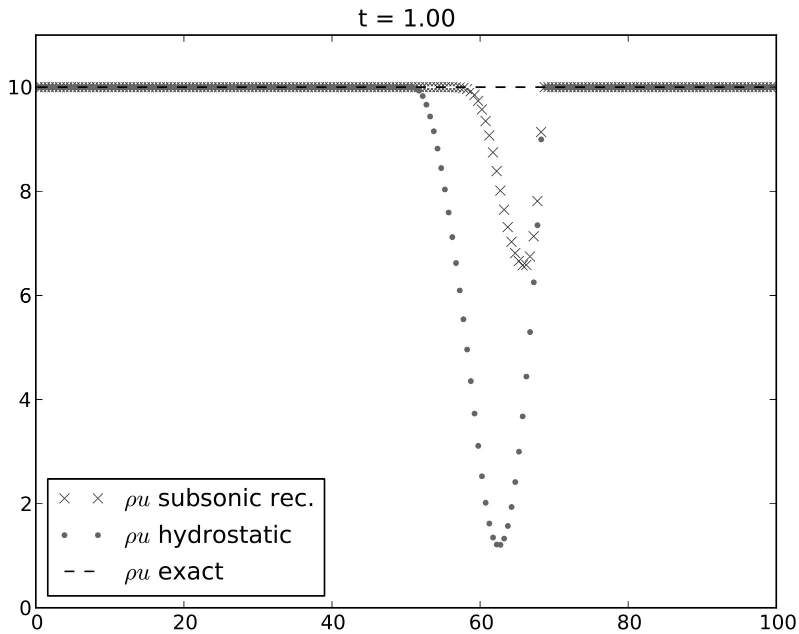

Fig. 5. *Supersonic steady state; ρu at $t = 1$.*

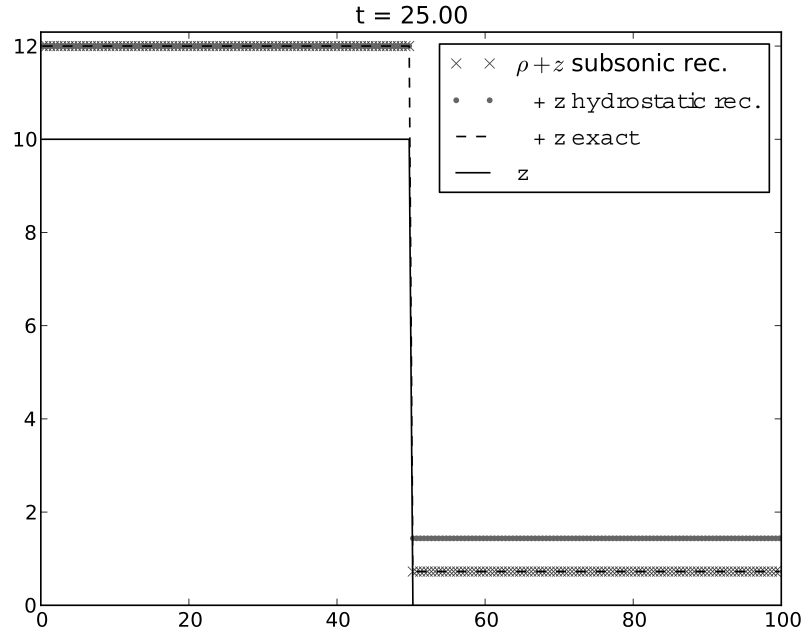

FIG. 6. *Supersonic steady state;* $\rho + z$ *at* $t = 25$.

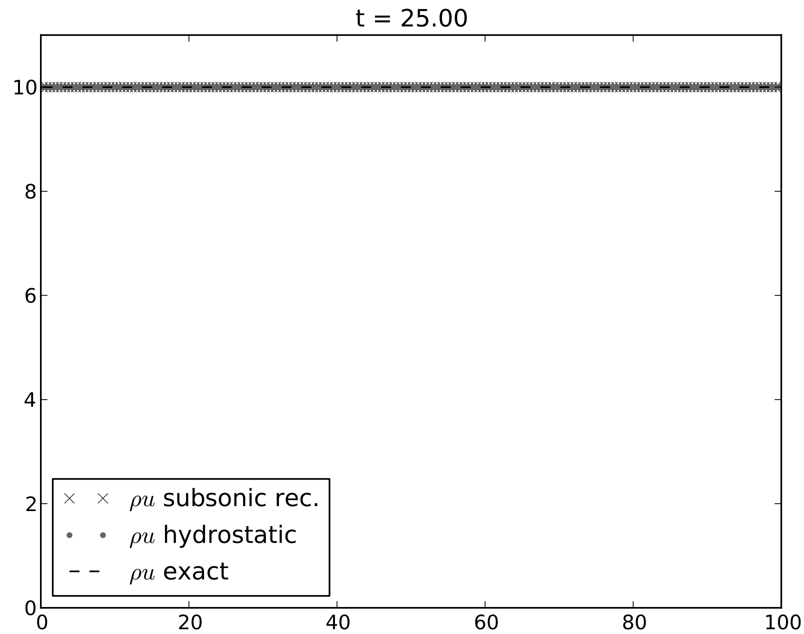

FIG. 7. *Supersonic steady state;* ρu *at* $t = 25$.

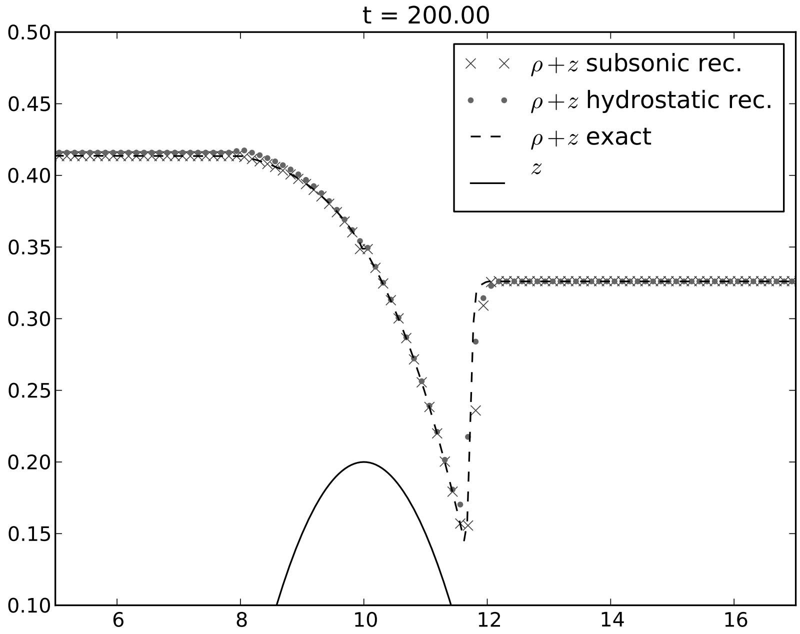

FIG. 8. *Transcritical flow with shock;* $\rho + z$ *at* $t = 200$.

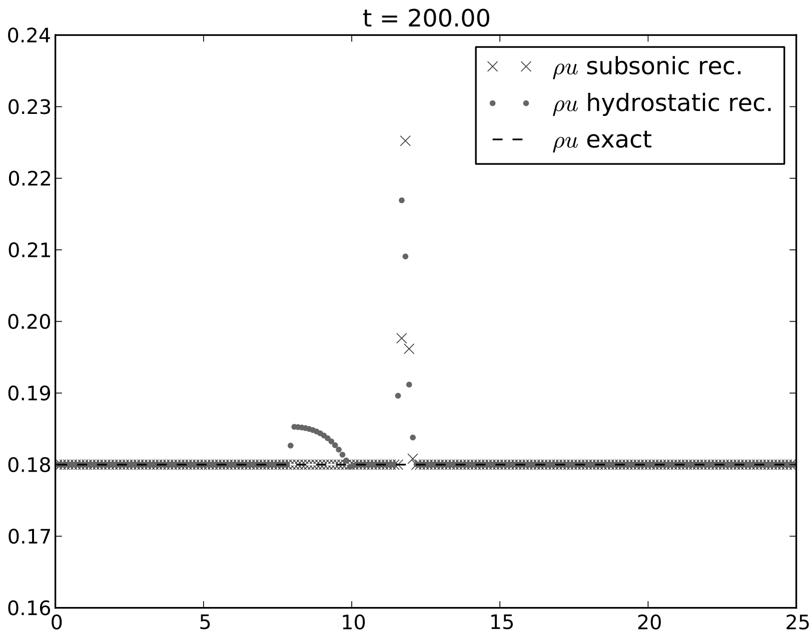

FIG. 9. *Transcritical flow with shock;* ρu *at* $t = 200$.

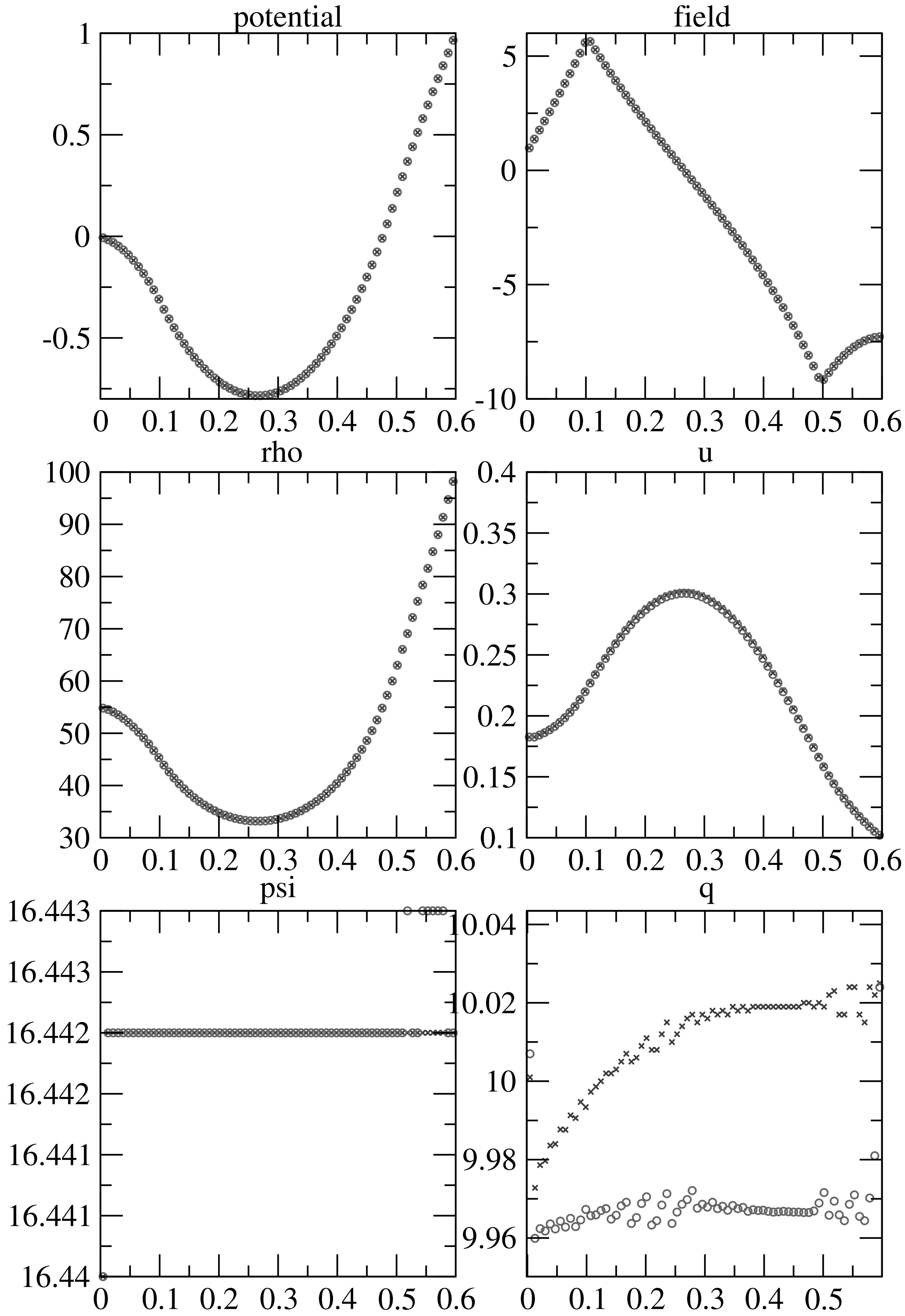

FIG. 10. *Crosses: subsonic reconstruction; circles: hydrostatic reconstruction. The quantity psi represents* $u^2/2 + e + p/\rho + \phi$*, and* $q = \rho u$*.*

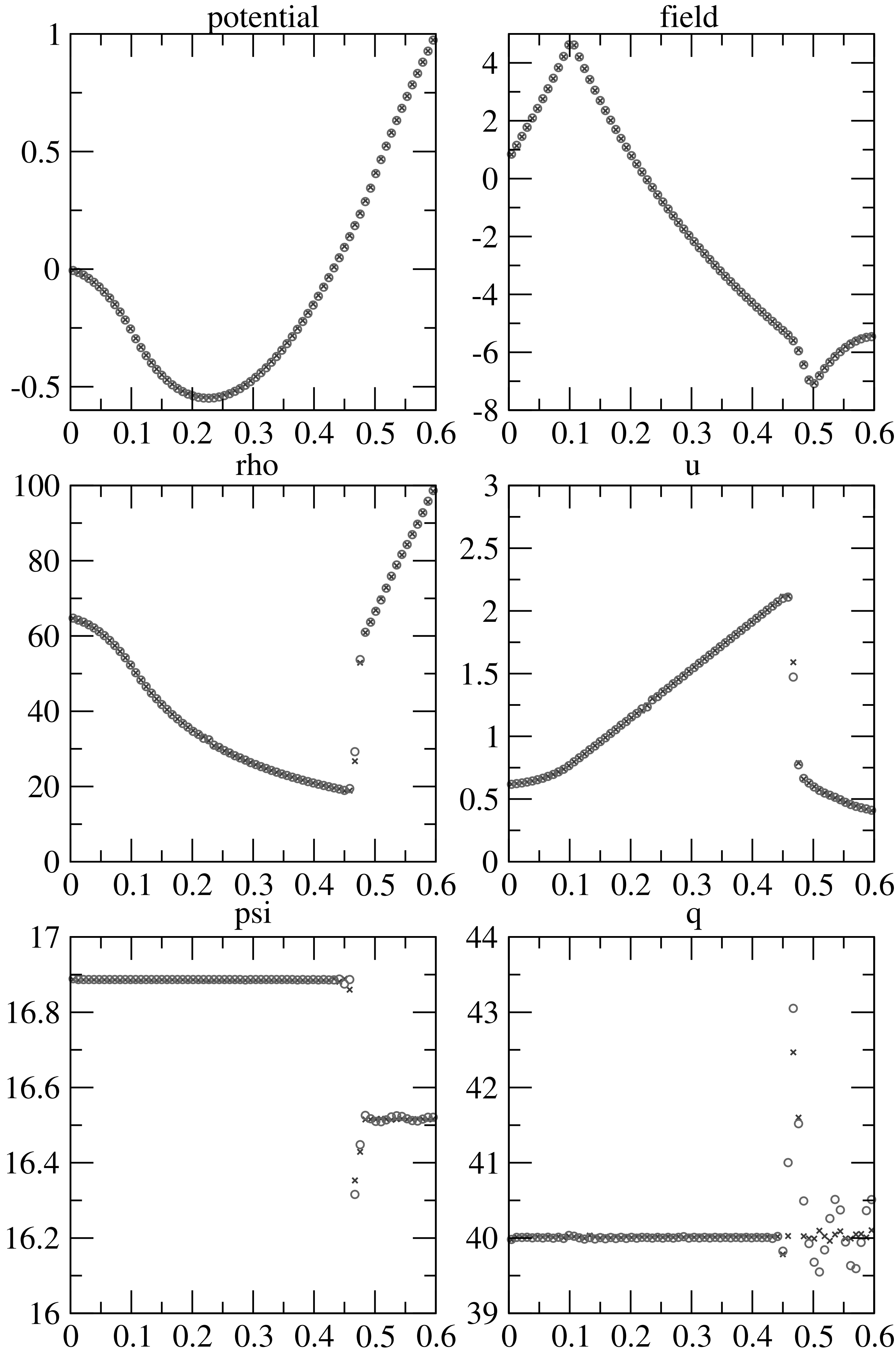

Fig. 11. *Crosses: subsonic reconstruction; circles: hydrostatic reconstruction. The quantity psi represents $u^2/2 + e + p/\rho + \phi$, and $q = \rho u$.*

In Figure 10 we show the result with the boundary discharge $q_0 = 10$. We use 100 points in space, and the final time is $t = 100$. A second-order reconstruction is used. As we see, we have reached a subsonic equilibrium where $u^2/2 + e + p/\rho + \phi$ is constant and q is almost constant.

Finally the same test with boundary discharge $q_0 = 40$ is shown in Figure 11. Even if we have reached a steady state, q and $u^2/2 + e + p/\rho + \phi$ are not constant. We observe a jump at the point where there is a change from a supersonic regime to a subsonic one. This is not surprising since the scheme is not exact for supersonic states.

REFERENCES

[1] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame, *A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows*, SIAM J. Sci. Comput., 25 (2004), pp. 2050–2065.

[2] E. Audusse, M.-O. Bristeau, and B. Perthame, *Kinetic Schemes for Saint-Venant Equations with Source Terms on Unstructured Grids*, INRIA report RR-3989, INRIA Roquencourt, Le Chesnay Cedex, France, 2000.

[3] D. S. Bale, R. J. LeVeque, S. Mitran, and J. A. Rossmanith, *A wave propagation method for conservation laws and balance laws with spatially varying flux functions*, SIAM J. Sci. Comput., 24 (2002), pp. 955–978.

[4] A. Bermúdez and M. E. Vásquez, *Upwind methods for hyperbolic conservation laws with source terms*, Comput. & Fluids, 23 (1994), pp. 1049–1071.

[5] R. Botchorishvili, B. Perthame, and A. Vasseur, *Equilibrium schemes for scalar conservation laws with stiff sources*, Math. Comp., 72 (2003), pp. 131–157.

[6] N. Botta, R. Klein, S. Langenberg, and S. Lützenkirchen, *Well balanced finite volume methods for nearly hydrostatic flows*, J. Comput. Phys., 196 (2004), pp. 539–565.

[7] F. Bouchut, *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws and Well-Balanced Schemes for Sources*, Front. Math., Birkhäuser Verlag, Basel, 2004.

[8] M. Castro, J. Macías, and C. Parés, *A Q-scheme for a class of systems of coupled conservation laws with source term, application to a two-layer 1-D shallow water system*, M2AN Math. Model. Numer. Anal., 35 (2001), pp. 107–127.

[9] M. Castro, A. Pardo Milanés, and C. Parés, *Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique*, Math. Models Methods Appl. Sci., 17 (2007), pp. 2055–2113.

[10] A. Chinnayya, A.-Y. LeRoux, and N. Seguin, *A well-balanced numerical scheme for the approximation of the shallow-water equations with topography: The resonance phenomenon*, Int. J. Finite Vol., 1 (2004), pp. 1–33.

[11] T. Gallouët, J.-M. Hérard, and N. Seguin, *Some approximate Godunov schemes to compute shallow-water equations with topography*, Comput. & Fluids, 32 (2003), pp. 479–513.

[12] L. Gosse, *A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms*, Comput. Math. Appl., 39 (2000), pp. 135–159.

[13] L. Gosse and A.-Y. Leroux, *Un schéma-équilibre adapté aux lois de conservation scalaires non-homogènes*, C. R. Acad. Sci. Paris Sér. I Math., 323 (1996), pp. 543–546.

[14] J. M. Greenberg and A. Y. Leroux, *A well-balanced scheme for the numerical processing of source terms in hyperbolic equations*, SIAM J. Numer. Anal., 33 (1996), pp. 1–16.

[15] S. Jin, *A steady-state capturing method for hyperbolic systems with geometrical source terms*, M2AN Math. Model. Numer. Anal., 35 (2001), pp. 631–645.

[16] A. Kurganov and D. Levy, *Central-upwind schemes for the Saint-Venant system*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 397–425.

[17] A. Kurganov and G. Petrova, *A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system*, Commun. Math. Sci., 5 (2007), pp. 133–160.

[18] S. Noelle, Y. Xing, and C. W. Shu, *High-order well-balanced finite volume WENO schemes for shallow water equation with moving water*, J. Comput. Phys., 226 (2007), pp. 29–58.

[19] S. Noelle, Y. Xing, and C. W. Shu, *High-order well-balanced schemes*, in Numerical Methods for Relaxation Systems and Balance Equations, Quad. Mat., Second University of Naples,

Naples, Italy, to appear.

[20] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*, Calcolo, 38 (2001), pp. 201–231.

[21] M. E. Vázquez-Cendón, *Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry*, J. Comput. Phys., 148 (1999), pp. 497–526.