Annals of Mathematics and Computer Science ISSN: 2789-7206
Vol 37 (2026) 1-31 https://doi.org/10.56947/amcs.v37.1000

NUMERICAL INSENSITIZING CONTROLS FOR AGE-STRUCTURED POPULATIONS

MIFIAMBA SOMA[1*], ALI OUEDRAOGO[2], YACOUBA ZONGO[3] and MAMADOU OUEDRAOGO[2]

Abstract. We compute controls which make an observation of an age-structured population with diffusion and external recruitment insensitive to small deformations of its habitat. A Hadamard formula gives the sensitivity as a boundary integral, which has a sign for nonnegative sources and is monotone in the data. Approximate controls are computed by penalized least squares, and exact controls for finitely many deformation modes by a Newton correction with a computable existence criterion. The scheme along characteristic lines is stable and the discrete controls converge. In one and two dimensions, the controls reduce the variation of the observation a hundredfold.

Keywords. Age-structured population dynamics, insensitizing controls, shape derivative, numerical approximation, conjugate gradient
2020 Mathematics Subject Classification. Primary 93-08, 49Q10; Secondary 65M06, 92D25

1. Introduction

We consider a population which lives in a habitat whose boundary is not well known. For example, the shoreline of a lake moves with the water level. A manager acts on a part ω of the habitat, by stocking or by removing individuals, and he follows the population in another part Θ. Since the shape of the habitat is uncertain, we want the observed quantity to depend on it as little as possible. In the framework of insensitizing controls introduced by J.-L. Lions (see [7]), this gives the following question. Can we choose the control so that the derivative of the observation with respect to the domain is zero, or as small as we want?

For the heat equation, this question was studied by Lissy, Privat and Simporé [19] and by Ervedoza, Lissy and Privat [12]. For age-structured population models with spatial diffusion and a renewal law, existence results were obtained in [23]. In that work, the shape derivative of the observation is a boundary integral of the normal derivatives of the state and of an adjoint state. Under a geometric

Date: Received: Oct 27, 2026; Revised: Nov 16, 2026; Accepted: Nov 21, 2026.
* Corresponding author
© The Author(s) 2026. This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of the licence, visit https://creativecommons.org/licenses/by-nc-nd/4.0/.

condition, there exist controls which make it as small as we want, and, in dimension one, controls which make it zero. These results are not constructive. As far as we know, no computation of insensitizing controls with respect to domain variations exists, even for the heat equation. The numerical works on insensitizing controls deal with unknown initial data [9]. The numerical approximation of null controls for parabolic equations has been studied a lot [14, 13, 8, 10, 18], and the control of age-structured models with diffusion is well developed [5, 20, 2, 3]. Optimal and time-optimal control problems for parabolic equations are studied in [4, 11], and the controllability of other evolution systems in [6].

In this paper, we compute such controls for a population with *external recruitment*, that is, without births in the habitat. Examples are sterile or stocked populations, or migratory species which reproduce elsewhere. This case corresponds to a zero fertility rate in [23]. Our problem has two particular features. First, the sensitivity is a quadratic function of the control, so we cannot make it zero by solving only one linear problem. Second, the model is a transport-diffusion equation in the time, age and space variables. Unlike reaction-diffusion models of populations such as Fisher's equation [1], it contains a transport term in age. So a one-dimensional habitat already gives a three-dimensional computational domain, and a two-dimensional habitat gives a four-dimensional one. Our contributions are as follows. In Section 3, we study the sensitivity in a qualitative way. For a nonnegative controlled source, the sensitivity is nonnegative, and in dimension one it vanishes only if the source vanishes. So, when the recruitment is not confined to ω, exact insensitizing needs controls for which the source $\xi + h\chi_\omega$ takes negative values. The sensitivity is also monotone with respect to the recruitment, the mortality and the observation region, homogeneous of degree two and invariant under the symmetries of the data. In Section 4, we compute approximate insensitizing controls by a penalized least-squares method and, for K deformation modes, exact insensitizing controls by a Newton correction in a space of dimension K, with an existence criterion that can be checked a posteriori. In Section 5, we discretize the model along the characteristic lines $t-a = \text{const}$, with exact discrete gradients. We prove the uniform stability of the discrete control-to-traces map, the consistency of the scheme, the convergence of the discrete penalized controls, and the convergence of the discrete exact controls under the criterion. Finally, in Sections 6 and 7, we validate the Hadamard formula, compute exact insensitizing controls, study the influence of the parameters in a quantitative way, and treat a two-dimensional habitat with four deformation modes.

2. The model and the shape derivative

Let $T > 0$ be the time horizon, $A > 0$ the maximal age and $U = (0, T) \times (0, A)$. The habitat $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2\}$, is an interval $(0, L)$ or a rectangle. The density

$y(t,a,x)$ of individuals of age a at time t and position x satisfies

$$\begin{cases} (\partial_t+\partial_a)y-\kappa\Delta y+\mu y=\xi+h\chi_\omega & \text{in } U\times\Omega,\\ y=0 & \text{on } U\times\partial\Omega,\\ y(t,0,x)=0,\quad y(0,a,x)=0, \end{cases} \tag{2.1}$$

where $\kappa>0$ is the diffusion coefficient ($\kappa=1$ if nothing else is said), $\mu=\mu(t,a)\geq 0$ is a bounded mortality rate, $\xi\in L^2(U\times\Omega)$ is the external recruitment, $\omega\subset\Omega$ is the control region and $h\in L^2(U\times\omega)$ is the control. The condition at age zero means that no individual is born in the habitat. Since ξ depends on a, it describes arrivals of individuals of all ages (migrants, restocking of fingerlings, release of sterile individuals). The observation is the L^2 norm of the population in a region Θ with $\overline{\Theta}\subset\Omega$,

$$J_h(\Omega)=\frac12\int_0^T\int_0^A\int_\Theta y^2\,dx\,da\,dt. \tag{2.2}$$

The data μ, ξ, ω, Θ, h are fixed in the physical space, and only the habitat is deformed. For a smooth vector field V and $|\tau|$ small, we set $\Omega_\tau=(\mathrm{Id}+\tau V)(\Omega)$. Let z be the solution of the backward adjoint problem

$$\begin{cases} -(\partial_t+\partial_a)z-\kappa\Delta z+\mu z=\chi_\Theta\, y & \text{in } U\times\Omega,\\ z=0 & \text{on } U\times\partial\Omega,\\ z(t,A,x)=0,\quad z(T,a,x)=0. \end{cases} \tag{2.3}$$

By the Hadamard formula of [23] (with zero fertility rate), the map $\tau\mapsto J_h(\Omega_\tau)$ is differentiable at $\tau=0$, and

$$\begin{aligned} \frac{d}{d\tau}J_h(\Omega_\tau)\Big|_{\tau=0}&=\kappa\int_{\partial\Omega}(V\cdot n)\,\Phi_h\,d\sigma,\\ \Phi_h(\sigma)&=\int_0^T\int_0^A\partial_n y(t,a,\sigma)\,\partial_n z(t,a,\sigma)\,da\,dt. \end{aligned} \tag{2.4}$$

(For $\kappa=1$, this is the formula of [23]. For a general κ, the factor κ is obtained by a scaling of the space variable.) In dimension one, $\partial\Omega=\{0,L\}$, $n(0)=-1$, $n(L)=1$, and the sensitivity reduces to the two numbers $\Phi_h(0)$ and $\Phi_h(L)$. In dimension two, Φ_h is a function on $\partial\Omega$, and in general we cannot cancel (2.4) for all fields V [23]. So we fix K deformation modes $V_1,\dots,V_K$ and we set

$$F_k(h)=\int_{\partial\Omega}(V_k\cdot n)\,\Phi_h\,d\sigma,\qquad k=1,\dots,K. \tag{2.5}$$

We say that h *exactly insensitizes* J for these modes if $F_k(h)=0$ for all k. In dimension one, we take $K=2$ and modes which move only one end point, so that $F=(\Phi_h(0),\Phi_h(L))$. For a rectangle, we take $K=4$ modes, and each of them translates one side. Then $F_k(h)$ is the integral of Φ_h over the k-th side.

Existence results. We write $y=y_\xi+y_h$, $z=z_\xi+z_h$, where (y_ξ,z_ξ) is the solution for $h=0$ and (y_h,z_h) is the solution for $\xi=0$. The map

$$\mathcal{L}:L^2(U\times\omega)\to H=L^2(U\times\partial\Omega)^2,\qquad \mathcal{L}h=\big(\partial_n y_h,\ \partial_n z_h\big)|_{U\times\partial\Omega},$$

is linear and bounded, and $h\mapsto\Phi_h$ is quadratic. In dimension one, let D_- and D_+ be the connected components of $\Omega\setminus\overline{\Theta}$ whose closures contain 0 and L. The

following facts are proved in [23]. (R1) If $\omega \cap \Theta \neq \emptyset$, or if ω meets D_- and D_+, the range of $\mathcal{L}$ is dense in H. (R2) Under the same assumptions, for every ξ, there exists a control which exactly insensitizes J. The proof of (R2) uses the Brouwer fixed point theorem and it is not constructive.

3. Qualitative properties of the sensitivity

Problems (2.1) and (2.3) only contain the operator $\partial_t + \partial_a$, so they decouple along the characteristic lines $t - a = s$. On each line, parametrized by $\sigma' \in (0, \ell)$, we have a forward and a backward heat equation in Ω, with zero data at the ends of the line. Let $p_r(x, x') > 0$ be the Dirichlet heat kernel of $\kappa\Delta$ in Ω, and let m, f be the restrictions of μ and of $f = \xi + h\chi_\omega$ to the line. Since μ does not depend on x, we have

$$\begin{aligned} y(\sigma', x) &= \int_0^{\sigma'} e^{-\int_s^{\sigma'} m} \int_\Omega p_{\sigma'-s}(x, x') f(s, x')\, dx' ds, \\ z(\sigma', x) &= \int_{\sigma'}^{\ell} e^{-\int_{\sigma'}^{s} m} \int_\Theta p_{s-\sigma'}(x, x') y(s, x')\, dx' ds, \end{aligned} \tag{3.1}$$

and the normal derivatives are given by the same formulas with p replaced by $\partial_{n,x} p$. For an interval or a rectangle, p is a product of one-dimensional kernels. By the Hopf lemma [22, Chap. 3], $-\partial_{n,x} p_r(\sigma, x') > 0$ for $r > 0$, $x' \in \Omega$ and σ in the relative interior of a side. The integrals converge in L^2 because $\|\partial_{n,x} p_r(\sigma, \cdot)\|_{L^2} \leq C r^{-3/4}$ in dimension one. In dimension two, the same estimate holds side by side after integration in the tangential variable.

We first give the sign of the sensitivity.

Proposition 3.1. *Assume that $f = \xi + h\chi_\omega \geq 0$ a.e. Then $y \geq 0$, $z \geq 0$, $\partial_n y \leq 0$, $\partial_n z \leq 0$, and $\Phi_h \geq 0$ on $\partial\Omega$. Moreover, in dimension one, for each $e \in \{0, L\}$, $\Phi_h(e) = 0$ if and only if $f = 0$ a.e. in $U \times \Omega$. In particular, if ξ does not vanish a.e. in $U \times (\Omega \setminus \omega)$, no control with $\xi + h\chi_\omega \geq 0$ can cancel one of the two sensitivities.*

Proof. The first assertions follow from (3.1) and from the positivity of p and $-\partial_n p$. In dimension one, on a given line, let σ_1 be the largest σ' such that $f = 0$ a.e. on $(0, \sigma') \times \Omega$. If $\sigma_1 < \ell$, then $y > 0$ on $(\sigma_1, \ell) \times (0, L)$ and $\partial_n y(\cdot, e) < 0$ a.e. on (σ_1, ℓ), because p and $-\partial_n p$ are positive. Hence the source $\chi_\Theta y$ of the adjoint problem is positive on $(\sigma_1, \ell) \times \Theta$. Since $z(\sigma')$ only depends on this source on (σ', ℓ), we get $z > 0$ and $\partial_n z(\cdot, e) < 0$ on $(0, \ell)$, in particular on (σ_1, ℓ). So $\partial_n y\, \partial_n z > 0$ there, and $\Phi_h(e) = 0$ implies $\sigma_1 = \ell$ on a.e. line, that is, $f = 0$ a.e. The converse is clear. □

The next result gives the monotonicity, the homogeneity and the symmetry of the sensitivity.

Proposition 3.2. *Let (f_i, μ_i, Θ_i), $i = 1, 2$, be two sets of data with $0 \leq f_1 \leq f_2$, $\mu_1 \geq \mu_2 \geq 0$ and $\Theta_1 \subset \Theta_2$, and let $\Phi^{(i)}$ be the corresponding sensitivities. Then:*

(i) *$0 \leq \Phi^{(1)} \leq \Phi^{(2)}$ on $\partial\Omega$. So the sensitivity increases with the controlled source and with the observation region, and it decreases when the mortality increases.*

(ii) *For every* $\lambda \in \mathbb{R}$, *if we replace* f *by* λf, *then* y *and* z *are multiplied by* λ *and* Φ *is multiplied by* λ^2.
(iii) *If* S *is an isometry of* $\mathbb{R}^d$ *with* $S(\Omega) = \Omega$, $S(\Theta) = \Theta$ *and* $f(t, a, Sx) = f(t, a, x)$, *then* $\Phi(S\sigma) = \Phi(\sigma)$ *on* $\partial\Omega$. *In dimension one, the reflection* $x \mapsto L - x$ *exchanges* $\Phi(0)$ *and* $\Phi(L)$.

Proof. (i) In (3.1), the kernels p and $-\partial_n p$ are positive, and the factors $e^{-\int m}$ are nonincreasing functions of the mortality. Hence $y_1 \le y_2$ and $0 \le -\partial_n y_1 \le -\partial_n y_2$. Then $0 \le \chi_{\Theta_1} y_1 \le \chi_{\Theta_2} y_2$ gives $0 \le -\partial_n z_1 \le -\partial_n z_2$. Finally, the product of two nonnegative nondecreasing quantities is nondecreasing. (ii) This follows from the linearity of (2.1)–(2.3). (iii) Since S preserves the Laplacian and the data, we have $y \circ S = y$ and $z \circ S = z$ by uniqueness, and S maps normal derivatives to normal derivatives. □

Propositions 3.1 and 3.2 have a simple interpretation. For a nonnegative population, that is, for every control which keeps the source nonnegative, a larger habitat never decreases the observation at first order, because the losses through the boundary are smaller. The sensitivity is larger when there are more arrivals, when the observation region is larger, and when the mortality is smaller. A control with $0 \le \xi + h\chi_\omega \le \xi$, that is, a control which only removes arrivals, can only reduce the sensitivity, by Proposition 3.2(i). In dimension one, a control which keeps the source nonnegative cannot cancel it. When the recruitment is not confined to ω, exact insensitizing needs controls for which the source, and therefore y, takes negative values. Such controls make sense when y is the deviation of the density from a reference population, as usual in control theory. Proposition 3.2 does not cover the monotonicity with respect to the diffusion coefficient. We study it numerically in Section 6.9.

4. Algorithms

4.1. **Penalized controls.** Let $b = (\partial_n y_\xi, \partial_n z_\xi)|_{U\times\partial\Omega} \in H$. Then the traces of the controlled solution are $\mathcal{L}h + b$. For $\varepsilon > 0$, we set

$$K_\varepsilon(h) = \frac{1}{2}\|h\|^2_{L^2(U\times\omega)} + \frac{1}{2\varepsilon}\|\mathcal{L}h + b\|^2_H. \tag{4.1}$$

Proposition 4.1. *For every* $\varepsilon > 0$, K_ε *has a unique minimizer* $h_\varepsilon = -\mathcal{L}^*(\varepsilon I + \mathcal{L}\mathcal{L}^*)^{-1}b$, *characterized by* $\varepsilon h_\varepsilon + \mathcal{L}^*(\mathcal{L}h_\varepsilon + b) = 0$. *If the range of* $\mathcal{L}$ *is dense, the residual* $r_\varepsilon = \mathcal{L}h_\varepsilon + b$ *tends to* 0 *in* H *as* $\varepsilon \to 0$, *and*

$$\|\Phi_{h_\varepsilon}\|_{L^1(\partial\Omega)} \le \tfrac{1}{2}\|r_\varepsilon\|^2_H, \qquad |F_k(h_\varepsilon)| \le \tfrac{1}{2}\|V_k \cdot n\|_{L^\infty(\partial\Omega)}\|r_\varepsilon\|^2_H \longrightarrow 0. \tag{4.2}$$

Moreover, (h_ε) *is bounded if and only if* b *belongs to the range of* $\mathcal{L}$. *In that case,* h_ε *converges to the control of minimal norm among the controls with* $\mathcal{L}h + b = 0$, *and* $\|r_\varepsilon\|^2_H = o(\varepsilon)$.

Proof. K_ε is continuous and strictly convex, with $K_\varepsilon'' = I + \varepsilon^{-1}\mathcal{L}^*\mathcal{L} \ge I$. This gives the minimizer and its characterization. By the spectral theorem, $r_\varepsilon = \varepsilon(\varepsilon I + \mathcal{L}\mathcal{L}^*)^{-1}b$ converges to the projection of b on $\ker \mathcal{L}^* = \{0\}$. The bound (4.2) follows from $|\partial_n y\, \partial_n z| \le \frac{1}{2}(|\partial_n y|^2 + |\partial_n z|^2)$, since $(\partial_n y, \partial_n z) = r_\varepsilon$. If $b = -\mathcal{L}\bar{h}$, then

$\frac{1}{2}\|h_\varepsilon\|^2 + \frac{1}{2\varepsilon}\|r_\varepsilon\|^2 = K_\varepsilon(h_\varepsilon) \le \frac{1}{2}\|\bar{h}\|^2$. A weak limit point h satisfies $\mathcal{L}h + b = 0$ and $\|h\| \le \|\bar{h}\|$ for every such $\bar{h}$, so it is the control of minimal norm $h^\dagger$. Then $\limsup \|h_\varepsilon\| \le \|h^\dagger\|$ gives the strong convergence and $\varepsilon^{-1}\|r_\varepsilon\|^2 \le \|h^\dagger\|^2 - \|h_\varepsilon\|^2 \to 0$. Conversely, if (h_ε) is bounded, a subsequence converges weakly to some h. Then $\mathcal{L}h_\varepsilon + b$ converges weakly to $\mathcal{L}h + b$, and also strongly to 0 since $r_\varepsilon \to 0$. So $\mathcal{L}h + b = 0$. □

4.2. **Exact correction.** Let h_0 be a control and let $g_1, \dots, g_K$ be directions in $L^2(U \times \omega)$. We set

$$G(c) = \Big(F_k(h(c))\Big)_{k=1}^K, \qquad h(c) = h_0 + \textstyle\sum_l c_l g_l, \qquad c \in \mathbb{R}^K.$$

Since y and z are affine in h, G is a polynomial of degree two. We write $G(c) = G(0) + Mc + R(c)$, with $M = DG(0)$ and $R_k(c) = c^\top S_k c$, where S_k is the symmetric matrix with entries $(S_k)_{lm} = \frac{1}{2}\int_{U\times\partial\Omega}(V_k \cdot n)\Big(\partial_n y_{g_l}\partial_n z_{g_m} + \partial_n y_{g_m}\partial_n z_{g_l}\Big)$. The gradient of F_k at h_0 is $\nabla F_k(h_0) = \mathcal{L}^* w_k$, with $w_k = \Big((V_k \cdot n)\partial_n z, (V_k \cdot n)\partial_n y\Big)$, where (y, z) is the solution associated with h_0. All these quantities only use the traces of the solutions associated with $h_0, g_1, \dots, g_K$, so they can be computed.

Proposition 4.2. *Let* $q = (\sum_k \|S_k\|^2)^{1/2}$, $\gamma = (\sum_l \|g_l\|^2)^{1/2}$. *Let* $\Gamma = 4\,q\,\|M^{-1}\|^2\,|G(0)|$. *Assume that* M *is invertible and that*

$$\Gamma = 4\,q\,\|M^{-1}\|^2\,|G(0)| < 1. \tag{4.3}$$

Then there is a unique $c \in \mathbb{R}^K$ *with* $|c| \le \rho$, *where* $\rho = 2\|M^{-1}\|\,|G(0)|$, *and* $G(c) = 0$. *The control* $h(c)$ *exactly insensitizes* J *for the modes* $V_1, \dots, V_K$, *and* $\|h(c) - h_0\| \le \gamma\rho$.

Proof. $G(c) = 0$ is equivalent to $c = \mathcal{T}(c)$, where $\mathcal{T}(c) = -M^{-1}(G(0) + R(c))$. We have $|R(c)| \le q|c|^2$ and $|R(c) - R(c')| \le q|c - c'|(|c| + |c'|)$. So, for $|c|, |c'| \le \rho$, we get $|\mathcal{T}(c)| \le \frac{\rho}{2}(1 + \Gamma) < \rho$ and $|\mathcal{T}(c) - \mathcal{T}(c')| \le \Gamma|c - c'|$. We conclude with the Banach fixed point theorem. □

This gives a two-stage algorithm. First, we compute $h_0 = h_\varepsilon$ for a moderate ε, so that $|G(0)|$ is small. We take $g_k = \nabla F_k(h_0)/\|\nabla F_k(h_0)\|$, so that M is invertible when the gradients are linearly independent. Then we solve $G(c) = 0$ by Newton's method from $c = 0$. Since G is an explicit quadratic map, each Newton iteration only uses the traces of the $K + 1$ solutions, which are computed once. Since $\|DG(c) - DG(c')\| \le 2q|c - c'|$, Newton's method converges under (4.3) by the Newton–Kantorovich theorem. The criterion (4.3) is computed before the correction. In our experiments, $|G(0)|$, and so Γ, decreases roughly like ε, while q and $\|M^{-1}\|$ change little. So we can obtain (4.3) by decreasing ε.

4.3. **The exact control of minimal norm.** Exact insensitizing controls are not unique. In general, the K conditions $F_k(h) = 0$ define a set of codimension K, and the control given by the two-stage algorithm depends on its starting point. We can choose the exact control of minimal norm, that is, a solution of

$$\min\Big\{\tfrac{1}{2}\|h\|^2 :\ h \in L^2(U \times \omega),\ F_k(h) = 0,\ k = 1, \dots, K\Big\}. \tag{4.4}$$

Since $h \mapsto F_k(h)$ is quadratic, we can write $F_k(h) = F_k(0) + (\ell_k, h) + \frac{1}{2}(\mathcal{A}_k h, h)$, where $\ell_k = \nabla F_k(0)$ and $\mathcal{A}_k$ is a bounded self-adjoint operator. Then $\nabla F_k(h) = \ell_k + \mathcal{A}_k h$. The Lagrange conditions for (4.4) are

$$h + \sum_{k=1}^{K} \lambda_k \nabla F_k(h) = 0, \qquad F_k(h) = 0, \quad k = 1, \dots, K, \tag{4.5}$$

for some multipliers $\lambda \in \mathbb{R}^K$.

Proposition 4.3. *Let $(h^\dagger, \lambda)$ satisfy (4.5), and assume that $I + \sum_k \lambda_k \mathcal{A}_k \geq \nu I$ for some $\nu > 0$. Then every exact insensitizing control h satisfies $\|h\|^2 \geq \|h^\dagger\|^2 + \nu\|h - h^\dagger\|^2$. In particular, $h^\dagger$ is the unique solution of (4.4).*

Proof. The Lagrangian $\mathcal{M}(h) = \frac{1}{2}\|h\|^2 + \sum_k \lambda_k F_k(h)$ is quadratic, with second derivative $I + \sum_k \lambda_k \mathcal{A}_k$. By (4.5), its gradient vanishes at $h^\dagger$. Hence $\mathcal{M}(h) \geq \mathcal{M}(h^\dagger) + \frac{\nu}{2}\|h - h^\dagger\|^2$. If h is an exact insensitizing control, $\mathcal{M}(h) = \frac{1}{2}\|h\|^2$, and $\mathcal{M}(h^\dagger) = \frac{1}{2}\|h^\dagger\|^2$. □

Remark 4.4. Proposition 4.3 does not give the existence of a solution of (4.4). The feasible set is nonempty by (R2), and minimizing sequences are bounded. But $\mathcal{L}$ is not compact. Since the problems decouple along the characteristic lines (Section 3), we have $\mathcal{L}\big(m(t-a)h\big) = m(t-a)\,\mathcal{L}h$ for every bounded function m of one variable. With $m_k(s) = \sqrt{2}\sin(ks)$, $h_k = m_k(t-a)h$ converges weakly to 0, but $\|\mathcal{L}h_k\|^2 \to \|\mathcal{L}h\|^2$. In the same way, $(\mathcal{A}_k h_k, h_k) \to (\mathcal{A}_k h, h)$, which is not zero in general. So the direct method of the calculus of variations does not apply. At the discrete level, the feasible set is closed and nonempty, and (4.4) has a solution. The open question is the convergence of these discrete solutions. The computations of Section 6.6 support it but do not prove it.

We compute $h^\dagger$ by a Lagrange–Newton method. For a given λ, the first equation of (4.5) is the linear system $\mathcal{H}(\lambda)h = -\sum_k \lambda_k \ell_k$, with $\mathcal{H}(\lambda) = I + \sum_k \lambda_k \mathcal{A}_k$ self-adjoint. We solve it by the MINRES method. Each application of $\sum_k \lambda_k \mathcal{A}_k$ costs two forward and two backward recursions along the characteristic lines. Then we solve the equations $F_k(h(\lambda)) = 0$ by Newton's method in λ. Differentiating $\mathcal{H}(\lambda)h(\lambda) = -\sum_k \lambda_k \ell_k$ gives $\partial_{\lambda_k} h = -\mathcal{H}(\lambda)^{-1}\nabla F_k(h(\lambda))$. So the Jacobian is the symmetric matrix $-\big(\mathcal{H}(\lambda)^{-1}\nabla F_k(h), \nabla F_l(h)\big)_{k,l}$. Under the assumption of Proposition 4.3, and if the gradients $\nabla F_k(h^\dagger)$ are linearly independent, this matrix is negative definite, and the method converges locally quadratically [21, Chap. 18]. Its global convergence is not guaranteed. For this reason, we start it from the two-stage control, and we initialize the multipliers by least squares. Another way to approximate $h^\dagger$ is to minimize $\frac{1}{2}\sum_k F_k(h)^2 + \frac{\eta}{2}\|h\|^2$ for a small $\eta > 0$ with the L-BFGS method, and then to apply the exact correction to the result.

5. Discretization and convergence

5.1. **The scheme.** We take the same step $\Delta = T/N_t$ in time and in age. Then the nodes $(t_n, a_j) = (n\Delta, j\Delta)$ and (t_{n+1}, a_{j+1}) are on the same characteristic

line, and the transport operator is discretized exactly along the lines, without numerical diffusion. In space, we use a uniform grid of step δx and the standard three-point (in dimension one) or five-point (in dimension two) approximation D of $-\kappa\Delta$ with Dirichlet conditions. For the deformed habitats used to check the Hadamard formula, the grid is stretched near the boundary, outside $\omega \cup \Theta$, and we use the three-point formula on nonuniform grids. With $\mathsf{M}_j = I + \Delta(\mathsf{D} + \mu(a_j))$, which is an M-matrix, the state and the adjoint state are computed by

$$\begin{aligned} y_j^0 = y_0^n = 0, \qquad \mathsf{M}_{j+1} y_{j+1}^{n+1} = y_j^n + \Delta\big(\xi_{j+1}^{n+1} + \chi_\omega h_{j+1}^{n+1}\big), \\ z_j^{N_t} = z_{N_a}^n = 0, \qquad \mathsf{M}_j z_j^n = z_{j+1}^{n+1} + \Delta \chi_\Theta y_j^n. \end{aligned} \tag{5.1}$$

Each time step needs N_a tridiagonal solves in dimension one, and N_a sparse solves in dimension two, with matrices factorized once. Since $\mathsf{M}_j^{-1} \geq 0$ entrywise, the scheme preserves nonnegativity. The discrete analogue of Proposition 3.1 holds when the normal derivatives are computed by the first-order formula $-u_1/\delta x$. We use second-order one-sided formulas, $(\partial_n u)(0) \approx -(4u_1 - u_2)/(2\delta x)$, and the discrete quantities

$$\Phi_h^\Delta(\sigma) = \Delta^2 \sum_{n,j} (\partial_n y)_j^n(\sigma)(\partial_n z)_j^n(\sigma), \qquad J_h^\Delta = \frac{\Delta^2}{2} \sum_{n,j} \sum_i |C_i| \chi_\Theta(x_i)(y_{j,i}^n)^2,$$
$$\|h\|^2 = \Delta^2 \sum_{n,j,i} |C_i| (h_{j,i}^n)^2,$$

where $|C_i|$ is the measure of the cell of the node x_i. In dimension two, the boundary integrals in F_k and in the norm of H use the boundary nodes with the weight δx. We follow the *discretize-then-optimize* approach. The gradients are the gradients of the discrete functions. They are computed with the transposes of (5.1), which are recursions along the same lines. For instance, the transpose of $f \mapsto y$ is $w \mapsto \Delta \mathsf{M}_j^{-\top} \lambda_j^n$ with $\lambda_j^n = w_j^n + \mathsf{M}_{j+1}^{-\top} \lambda_{j+1}^{n+1}$. We checked each transpose by the dot-product test, with relative errors of order 10^{-15}. The penalized controls are computed by the conjugate gradient method applied to the discrete normal equations. Iterative methods of this kind for least-squares problems are also used for identification problems [16]. Each iteration costs two forward and two backward recursions.

5.2. **Stability and convergence.** We give the analysis in dimension one. Let $\mathcal{L}_\Delta$ be the discrete control-to-traces map and $\|g\|_{H_\Delta}^2 = \Delta^2 \sum_{n,j} |g_j^n|^2$.

Lemma 5.1. *Assume that the spatial grid is uniform, with step $\delta x = L/(N_x + 1)$ and $N_x \geq 2$, that $\mu = \mu(a) \geq 0$ and that $\kappa = 1$. Then, for the first-order and for the second-order formulas for the normal derivatives,*

$$\|\mathcal{L}_\Delta h\|_{H_\Delta}^2 \leq 10L\Big(1 + \frac{L^4}{16}\Big)\|h\|^2$$

for every discrete control h, independently of Δ and δx.

Proof. For $u \in \mathbb{R}^{N_x}$, extended by $u_0 = u_{N_x+1} = 0$, let $g_{i+1/2} = (u_{i+1} - u_i)/\delta x$, $0 \leq i \leq N_x$, and let $\|u\|^2 = \delta x \sum_i u_i^2$ and $\|g\|^2 = \delta x \sum_i g_{i+1/2}^2$. Summation by parts gives $(\mathsf{D}u, u) = \|g\|^2$, where $(\cdot, \cdot)$ is the inner product associated with

$\|\cdot\|$. The smallest eigenvalue of D is $\lambda_1 = 4\delta x^{-2}\sin^2(\pi\delta x/(2L)) \geq 4/L^2$, because $\sin\theta \geq 2\theta/\pi$ on $[0,\pi/2]$. Then $\|g\|^2 = (\mathsf{D}u,u) \leq \|\mathsf{D}u\|\,\|u\|$ and $\lambda_1\|u\|^2 \leq \|g\|^2$ give

$$\|u\| \leq \frac{L^2}{4}\|\mathsf{D}u\|, \qquad \|g\| \leq \frac{L}{2}\|\mathsf{D}u\|. \tag{5.2}$$

Discrete trace inequality. Since $(\mathsf{D}u)_i = -(g_{i+1/2} - g_{i-1/2})/\delta x$, we have, for all k, m,

$$|g_{k+1/2}^2 - g_{m+1/2}^2| \leq \sum_{i=1}^{N_x} \delta x\,|(\mathsf{D}u)_i|\,|g_{i+1/2} + g_{i-1/2}| \leq 2\|\mathsf{D}u\|\,\|g\|.$$

We choose m such that $g_{m+1/2}^2 \leq \|g\|^2/L$ (recall that $(N_x+1)\delta x = L$). With (5.2), we get $\max_k g_{k+1/2}^2 \leq \|g\|^2/L + 2\|g\|\,\|\mathsf{D}u\| \leq \frac{5L}{4}\|\mathsf{D}u\|^2$. The first-order formula at $x = 0$ is $-u_1/\delta x = -g_{1/2}$, and the second-order one is $-(4u_1 - u_2)/(2\delta x) = -(3g_{1/2} - g_{3/2})/2$. It is similar at $x = L$. In both cases, $|(\partial_n u)(e)| \leq 2\max_k |g_{k+1/2}|$, and so

$$|(\partial_n u)(0)|^2 + |(\partial_n u)(L)|^2 \leq 10L\,\|\mathsf{D}u\|^2. \tag{5.3}$$

Energy estimate along the lines. Let $u = y_{j+1}^{n+1}$, $v = y_j^n$ and $f = \chi_\omega h_{j+1}^{n+1}$. By (5.1), $u - v + \Delta(\mathsf{D} + \mu(a_{j+1}))u = \Delta f$. We take the inner product with $\mathsf{D}u$. We use $(u-v,\mathsf{D}u) = \frac{1}{2}\big(\|g_u\|^2 - \|g_v\|^2 + \|g_u - g_v\|^2\big)$, $\mu(\mathsf{D}u,u) \geq 0$ and $\Delta(f,\mathsf{D}u) \leq \frac{\Delta}{2}\|f\|^2 + \frac{\Delta}{2}\|\mathsf{D}u\|^2$, and we obtain

$$\|g_u\|^2 - \|g_v\|^2 + \Delta\|\mathsf{D}u\|^2 \leq \Delta\|f\|^2.$$

On each characteristic line, y vanishes at the first node. We sum along each line, and then over the lines, and we get $\Delta^2\sum_{n,j}\|\mathsf{D}y_j^n\|^2 \leq \Delta^2\sum_{n,j}\|\chi_\omega h_j^n\|^2 \leq \|h\|^2$. The same computation for the backward recursion of (5.1), with the source $\chi_\Theta y$, and (5.2) give

$$\Delta^2\sum_{n,j}\|\mathsf{D}z_j^n\|^2 \leq \Delta^2\sum_{n,j}\|y_j^n\|^2 \leq \frac{L^4}{16}\,\Delta^2\sum_{n,j}\|\mathsf{D}y_j^n\|^2 \leq \frac{L^4}{16}\|h\|^2.$$

The conclusion follows from (5.3), applied to y_j^n and z_j^n at every node. □

The proof gives more than the statement. For any source f in (5.1) (in place of $\xi + \chi_\omega h$), the solution satisfies

$$\begin{gathered}\Delta^2\sum_{n,j}\|\mathsf{D}y_j^n\|^2 \leq \|f\|_\Delta^2, \qquad \|y\|_\Delta \leq \frac{L^2}{4}\|f\|_\Delta,\\ \Delta^2\sum_{n,j}\sum_e |(\partial_n y)_j^n(e)|^2 \leq 10L\|f\|_\Delta^2,\end{gathered} \tag{5.4}$$

where $\|f\|_\Delta^2 = \Delta^2\sum_{n,j}\|f_j^n\|^2$. The same estimates hold for the backward recursion of (5.1) with any source. The constant of Lemma 5.1 is not sharp. For the data of Section 6, the computed norm of $\mathcal{L}_\Delta$ is 0.35 for $N = 20$ and $N = 40$. All the controls are computed on the uniform grid. The deformed grids used in Section 6.2 are only used to evaluate the observation.

5.3. **Discrete and continuous spaces.** We now compare the discrete controls with the continuous ones. Let $X = L^2(U \times \omega)$ and $H = L^2(U)^4$ (the space H of Section 2 for $d = 1$), and let X_Δ, H_Δ be the corresponding discrete spaces, with the weighted inner products defined above. We denote by $\mathsf{P}_\Delta : X_\Delta \to X$ and $\mathsf{E}_\Delta : H_\Delta \to H$ the extensions by piecewise constant functions. They are isometries onto L^2 of the union of the cells of the nodes. This union differs from $U \times \omega$, resp. U, by a set of measure $O(\Delta + \delta x)$. We denote by $\mathsf{R}_\Delta = \mathsf{P}_\Delta^*$, $\mathsf{Q}_\Delta = \mathsf{E}_\Delta^*$ the corresponding averaging operators, so that $\mathsf{P}_\Delta\mathsf{R}_\Delta \to I$ and $\mathsf{E}_\Delta\mathsf{Q}_\Delta \to I$ strongly. Let b_Δ be the discrete traces of the uncontrolled solution, and h_ε^Δ the minimizer of the discrete functional K_ε^Δ. We consider the limit $\Delta \to 0$, $\delta x \to 0$ (independently, with A/Δ an integer), under the following assumption:

(C) the spatial grid is uniform, $\kappa = 1$, $\mu = \mu(a)$ with $\mu \in C^1([0, A])$ and $\mu \geq 0$, and $\xi \in C(\overline{U} \times [0, L])$.

This assumption holds in all our experiments. We use the following three properties of the scheme:

(A1) $\sup_\Delta \|\mathcal{L}_\Delta\| < \infty$;
(A2) $\mathsf{E}_\Delta\mathcal{L}_\Delta\mathsf{R}_\Delta h \to \mathcal{L}h$ in H for every h in a dense subset of X, and $\mathsf{E}_\Delta b_\Delta \to b$ in H;
(A3) $\mathsf{P}_\Delta\mathcal{L}_\Delta^*\mathsf{Q}_\Delta g \to \mathcal{L}^* g$ in X for every g in a dense subset of H.

Property (A1) is the stability of the scheme for the traces. It holds under (C) by Lemma 5.1. Property (A2) is the convergence of the scheme for the traces of the state and of the adjoint state, and (A3) is the same property for the discrete transposed recursions. We prove (A2) and (A3) under (C) in the next subsection (Proposition 5.5). By (A1), the convergences in (A2) and (A3) then hold for every $h \in X$ and every $g \in H$, and they are uniform on compact sets. Moreover, if $h_\Delta \in X_\Delta$ and $\mathsf{P}_\Delta h_\Delta \to h$ in X, then $\mathsf{E}_\Delta\mathcal{L}_\Delta h_\Delta \to \mathcal{L}h$ in H, because $\|\mathsf{E}_\Delta\mathcal{L}_\Delta(h_\Delta - \mathsf{R}_\Delta h)\| \leq \sup_\Delta \|\mathcal{L}_\Delta\| \, \|\mathsf{P}_\Delta h_\Delta - \mathsf{P}_\Delta\mathsf{R}_\Delta h\|$. The same holds for $\mathcal{L}_\Delta^*$.

5.4. **Consistency of the scheme.** We now prove (A2) and (A3). The proof uses the characteristic structure. On every characteristic line, each continuous problem is a one-dimensional heat equation, and the scheme is the implicit Euler scheme for this equation. For a line, we denote by σ the parameter along the line, so that $\partial_\sigma = \partial_t + \partial_a$. Let $\mathcal{N} \subset (0, L)$ be a finite set. We say that a function u on $\overline{U} \times [0, L]$ has *regularity* $(\mathcal{R}_\mathcal{N})$ if u and $\partial_x u$ are continuous on $\overline{U} \times [0, L]$, u is Lipschitz continuous along the lines, $\partial_\sigma u$ is bounded, $\partial_x u$ is Lipschitz continuous with respect to x (so $\partial_{xx} u$ is bounded), and $\partial_\sigma u$ and $\partial_{xx} u$ are continuous on $\overline{U} \times ([0, L] \setminus \mathcal{N})$. For such a function, we denote by $\hat{u}$ its values at the nodes, including the boundary nodes $x_0 = 0$ and $x_{N_x+1} = L$, and by $\overline{\mathsf{D}}$ the three-point operator D applied with these boundary values. The discrete norm $\|\cdot\|_\Delta$ is the norm of (5.4).

Lemma 5.2. *Let u have regularity $(\mathcal{R}_\mathcal{N})$. Let $\zeta_j^n = (\hat{u}_j^n - \hat{u}_{j-1}^{n-1})/\Delta - \widehat{\partial_\sigma u}_j^n$ for $n, j \geq 1$, $\zeta_j'^n = (\hat{u}_{j+1}^{n+1} - \hat{u}_j^n)/\Delta - \widehat{\partial_\sigma u}_j^n$ for $n < N_t$, $j < N_a$, and $\varrho_j^n = \overline{\mathsf{D}}\hat{u}_j^n + \widehat{\partial_{xx} u}_j^n$, with the convention $\partial_{xx} u = 0$ on $\overline{U} \times \mathcal{N}$. Then $\|\zeta\|_\Delta + \|\zeta'\|_\Delta + \|\varrho\|_\Delta \to 0$ as $\Delta, \delta x \to 0$.*

Proof. Let M be a bound of $|\partial_\sigma u|$ and $|\partial_{xx} u|$. Since u is Lipschitz continuous along the lines, $\zeta^n_{j,i}$ is the mean value of $\partial_\sigma u(\cdot, x_i) - \partial_\sigma u(t_n, a_j, x_i)$ on the segment of the line between (t_{n-1}, a_{j-1}) and (t_n, a_j). In the same way, since $\partial_x u$ is Lipschitz continuous, $-\overline{\mathsf{D}}\hat{u}^n_{j,i}$ is a weighted mean value of $\partial_{xx} u(t_n, a_j, \cdot)$ on (x_{i-1}, x_{i+1}), with the weight $(1 - |x - x_i|/\delta x)/\delta x$. Hence $|\zeta|, |\zeta'|, |\varrho| \le 2M$ at every node. Let $\eta > 0$, and let $\mathcal{N}_\eta$ be the set of the points at distance less than η from $\mathcal{N}$. On the compact set $\overline{U} \times ([0, L] \setminus \mathcal{N}_\eta)$, $\partial_\sigma u$ and $\partial_{xx} u$ are uniformly continuous. So, if Δ and δx are small enough, $|\zeta| + |\zeta'| + |\varrho| \le \eta$ at all the nodes $x_i \notin \mathcal{N}_{2\eta}$. The nodes $x_i \in \mathcal{N}_{2\eta}$ have a total weight $\delta x \, \#\{i : \ x_i \in \mathcal{N}_{2\eta}\} \le (4\eta + 2\delta x) \, \#\mathcal{N}$. Therefore $\limsup(\|\zeta\|^2_\Delta + \|\zeta'\|^2_\Delta + \|\varrho\|^2_\Delta) \le 3\Big(TAL\,\eta^2 + 4M^2 TA\,4\eta\,\#\mathcal{N}\Big)$, and η is arbitrary. □

Lemma 5.3. *Assume* (C).

(i) *If $f \in C^\infty_c(U \times (0, L))$, the solution v of $\partial_\sigma v - \partial_{xx} v + \mu v = f$ with $v = 0$ at $x \in \{0, L\}$ and at the beginning of the lines, and the solution of the backward problem $-\partial_\sigma v - \partial_{xx} v + \mu v = f$ with $v = 0$ at $x \in \{0, L\}$ and at the end of the lines, have regularity $(\mathcal{R}_\emptyset)$, and vanish near the beginning, resp. the end, of the lines.*

(ii) *The same holds for $f = 0$ and the boundary conditions $v(\cdot, \cdot, 0) = g_0$, $v(\cdot, \cdot, L) = g_L$, with $g_0, g_L \in C^\infty_c(U)$.*

(iii) *If v is a forward solution as in* (i) *or* (ii)*, then the solution w of $-\partial_\sigma w - \partial_{xx} w + \mu w = \chi_\Theta v$, with $w = 0$ at $x \in \{0, L\}$ and at the end of the lines, has regularity $(\mathcal{R}_{\partial\Theta})$.*

Proof. On each line, these are one-dimensional heat equations with the potential $\mu(a)$, which is of class C^1 along the line. In (i) and (ii), the data are smooth and vanish near the corners of the domain $(0, \ell) \times (0, L)$ of the line. So the compatibility conditions hold at all orders, and the solutions are smooth up to the boundary [17, Chap. IV]. The bounds are uniform with respect to the line, because the data are smooth functions of (t, a). The solutions vanish near the beginning (resp. the end) of the lines because the data do. In (iii), w vanishes on $\{x = 0\}$ and $\{x = L\}$, so $\partial_\sigma w$ also vanishes there. The function $w_1 = \partial_\sigma w$ solves $-\partial_\sigma w_1 - \partial_{xx} w_1 + \mu w_1 = \chi_\Theta \partial_\sigma v - \mu' w$, with the final value $w_1 = -\chi_\Theta v$ at the end of the lines and with $w_1 = 0$ at $x \in \{0, L\}$. The final value follows from the equation at the end of the lines, where $w = \partial_{xx} w = 0$. These data are compatible at the corners, since $\overline{\Theta} \subset (0, L)$. This is justified by differentiating the Duhamel formula. The source and the final value are bounded, so w_1 is bounded by the maximum principle. It is continuous except at the points of discontinuity of its final value, which lie on $\overline{U} \times \partial\Theta$ [17, Chap. IV]. Since the source $\chi_\Theta v$ is bounded, $w \in W^{2,1}_p$ on each line for every $p < \infty$ [17, Chap. IV], uniformly with respect to the line. So w and $\partial_x w$ are Hölder continuous. Their continuity with respect to (t, a) follows from the continuous dependence of the data on the line. Then $\partial_{xx} w = -\partial_\sigma w + \mu w - \chi_\Theta v$ is bounded, and continuous outside $\overline{U} \times \partial\Theta$, and $\partial_x w$ is Lipschitz continuous with respect to x. □

The second tool is a discrete Green formula. We write $(\cdot, \cdot)_\Delta$ for the inner product associated with $\| \cdot \|_\Delta$, and $(\cdot, \cdot)_{H_\Delta}$ for the inner product of H_Δ. We

denote by $\partial_n^{(1)}$ the first-order normal derivative $-u_1/\delta x$ at $x = 0$ and $-u_{N_x}/\delta x$ at $x = L$.

Lemma 5.4. *Let y be the solution of (5.1) with a source f in place of $\xi+\chi_\omega h$, with $f_j^n = 0$ if $n = 0$ or $j = 0$ (these values are not used by (5.1)). Let $\Psi_j^n \in \mathbb{R}^{N_x+2}$, $0 \le n \le N_t$, $0 \le j \le N_a$, with $\Psi_j^n = 0$ if $n = N_t$ or $j = N_a$. Then*

$$(f, \Psi)_\Delta = (y, \mathsf{S}^*\Psi)_\Delta - \Delta^2 \sum_{n,j} \sum_{e\in\{0,L\}} (\partial_n^{(1)} y)_j^n(e)\, \Psi_j^n(e),$$

$$\mathsf{S}^*\Psi_j^n = \frac{\Psi_j^n - \Psi_{j+1}^{n+1}}{\Delta} + \overline{\mathsf{D}}\Psi_j^n + \mu(a_j)\Psi_j^n,$$

where $\Psi_j^n(0)$ and $\Psi_j^n(L)$ are the boundary values of Ψ_j^n, and $\Psi_{j+1}^{n+1} = 0$ if $n = N_t$ or $j = N_a$. The same formula holds for the solution z of the backward recursion of (5.1) with a source r such that $r_j^n = 0$ if $n = N_t$ or $j = N_a$, with $\Psi_j^n = 0$ if $n = 0$ or $j = 0$, and with Ψ_{j-1}^{n-1} in place of Ψ_{j+1}^{n+1}.

Proof. By (5.1), $f_j^n = (y_j^n - y_{j-1}^{n-1})/\Delta + \mathsf{D}y_j^n + \mu(a_j)y_j^n$ for $n, j \ge 1$. Since y vanishes for $n = 0$ or $j = 0$, a change of index gives $\sum_{n,j\ge1}(y_{j-1}^{n-1}, \Psi_j^n) = \sum_{n,j\ge1}(y_j^n, \Psi_{j+1}^{n+1})$, where $(\cdot,\cdot)$ is the inner product of $\mathbb{R}^{N_x}$ weighted by δx. In space, summation by parts gives $(\mathsf{D}y, \Psi) = (y, \overline{\mathsf{D}}\Psi) + (y_1\Psi(0) + y_{N_x}\Psi(L))/\delta x$, and $y_1/\delta x = -(\partial_n^{(1)}y)(0)$, $y_{N_x}/\delta x = -(\partial_n^{(1)}y)(L)$. The proof for the backward recursion is the same. □

Proposition 5.5. *Under assumption (C), properties (A2) and (A3) hold, for the first-order and for the second-order formulas for the normal derivatives.*

Proof. (A2). Let $h \in C_c^\infty(U \times \omega)$, which is a dense subset of X. Let y be the continuous state associated with h (with $\xi = 0$) and z the corresponding adjoint state. By Lemma 5.3, y has regularity $(\mathcal{R}_\emptyset)$ and z has regularity $(\mathcal{R}_{\partial\Theta})$. Since $\mathsf{R}_\Delta h - \hat{h} \to 0$ uniformly and $\mathcal{L}_\Delta$ is bounded uniformly, we can replace $\mathsf{R}_\Delta h$ by $\hat{h}$. Let y_Δ, z_Δ be the discrete solutions associated with $\hat{h}$. The errors $\varphi_y = y_\Delta - \hat{y}$ and $\varphi_z = z_\Delta - \hat{z}$ vanish at the beginning, resp. the end, of the lines. They solve the forward, resp. backward, recursion of (5.1) with the sources $-\upsilon_y$ and $\chi_\Theta\varphi_y - \upsilon_z$, where υ_y and υ_z are the residuals of the schemes applied to $\hat{y}$ and $\hat{z}$. At the nodes $x_i \notin \partial\Theta$, the continuous equations give $\upsilon_y = \zeta_y + \varrho_y$ and $\upsilon_z = -\zeta_z' + \varrho_z$, with the notation of Lemma 5.2. At the nodes $x_i \in \partial\Theta$, υ_z is bounded, and these nodes have a total weight at most $\delta x\, \#\partial\Theta$. Hence $\|\upsilon_y\|_\Delta + \|\upsilon_z\|_\Delta \to 0$ by Lemma 5.2, and, by (5.4),

$$\Delta^2 \sum_{n,j} \sum_e \Big(|(\partial_n\varphi_y)_j^n(e)|^2 + |(\partial_n\varphi_z)_j^n(e)|^2 \Big) \le C\Big(\|\upsilon_y\|_\Delta^2 + \|\upsilon_z\|_\Delta^2 \Big) \to 0.$$

The functions y and z are of class C^2 in x near $x = 0$ and $x = L$, uniformly in (t, a). So the one-sided formulas applied to $\hat{y}$ and $\hat{z}$ converge uniformly to $\partial_n y$ and $\partial_n z$ at the nodes. The piecewise constant functions built on the nodal values of these continuous functions converge in $L^2(U)$. Hence $\mathsf{E}_\Delta\mathcal{L}_\Delta\hat{h} \to \mathcal{L}h$. For b_Δ, let $\xi_k \in C_c^\infty(U \times (0, L))$ converge to ξ in L^2. The previous argument, applied with the source ξ_k in the whole interval, gives $\mathsf{E}_\Delta b_\Delta(\xi_k) \to b(\xi_k)$. By (5.4), applied to the state and to the adjoint state, $\|b_\Delta(\xi) - b_\Delta(\xi_k)\|_{H_\Delta} \le C\|\hat{\xi} - \hat{\xi}_k\|_\Delta$. This tends

to $C\|\xi-\xi_k\|_{L^2}$ as $\Delta,\delta x\to 0$, because $\xi-\xi_k$ is continuous. By the continuous version of (5.4), $\|b(\xi)-b(\xi_k)\|_H\le C\|\xi-\xi_k\|_{L^2}$. Letting $k\to\infty$, we get $\mathsf{E}_\Delta b_\Delta\to b$.

(A3). Let $g=(g_0^y,g_L^y,g_0^z,g_L^z)\in C_c^\infty(U)^4$, which is a dense subset of H. The components correspond to the traces $\partial_n y(e)$ and $\partial_n z(e)$. Let Ψ be the backward solution of Lemma 5.3(ii) with the boundary values $-g_e^y$, Υ the forward solution with the boundary values $-g_e^z$, and ψ the backward solution of Lemma 5.3(iii) with the source $\chi_\Theta\Upsilon$. Integrations by parts along the lines show that $(\mathcal{L}h,g)_H=(h,\Psi+\psi)_X$ for every $h\in X$, that is, $\mathcal{L}^*g=(\Psi+\psi)|_{U\times\omega}$. In the same way, $\mathcal{L}_\Delta^*\hat g=\chi_\omega(\Psi_\Delta+\psi_\Delta)$. Here Ψ_Δ is the transpose of $f\mapsto\partial_n y$ (with y the solution of (5.1) with source f) applied to $(\hat g_0^y,\hat g_L^y)$, Υ_Δ is the transpose of $r\mapsto\partial_n z$ applied to $(\hat g_0^z,\hat g_L^z)$, and ψ_Δ is the transpose of $f\mapsto y$ applied to $\chi_\Theta\Upsilon_\Delta$ (restricted or extended by zero to the nodes of (5.1), which only changes one node per line).

Let f be arbitrary and y the solution of (5.1) with source f. By definition of Ψ_Δ, $(f,\Psi_\Delta)_\Delta=(\partial_n y,\hat g^y)_{H_\Delta}$. Lemma 5.4, applied to $\hat\Psi$ (whose boundary values are $-\hat g_e^y$ and which vanishes at the end of the lines), gives $(f,\hat\Psi)_\Delta=(y,\mathsf{S}^*\hat\Psi)_\Delta+(\partial_n^{(1)}y,\hat g^y)_{H_\Delta}$. Since Ψ solves the homogeneous equation, $\mathsf{S}^*\hat\Psi=-\zeta'_\Psi+\varrho_\Psi$ at the nodes with $n<N_t$ and $j<N_a$, with the notation of Lemma 5.2. At the last node of each line, $\hat\Psi=0$ and $\mathsf{S}^*\hat\Psi=0$. With the second-order formula, $(\partial_n y-\partial_n^{(1)}y)(0)=-(2y_1-y_2)/(2\delta x)=-\frac{\delta x}{2}(\mathsf{D}y)_1$ and $|(\mathsf{D}y)_1|\le\delta x^{-1/2}\|\mathsf{D}y\|$. It is similar at $x=L$. With the first-order formula, this difference is zero. Hence, by (5.4),

$$|(f,\Psi_\Delta-\hat\Psi)_\Delta|\le\Big(\sqrt{\tfrac{\delta x}{2}}\,\|\hat g^y\|_{H_\Delta}+\tfrac{L^2}{4}\|\mathsf{S}^*\hat\Psi\|_\Delta\Big)\|f\|_\Delta,$$

and $\|\Psi_\Delta-\hat\Psi\|_\Delta\to 0$ by Lemma 5.2. (The inequality holds for f supported on the nodes with $n,j\ge 1$. At the first node of each line, $\Psi_\Delta=0$, and these fewer than N_t+N_a+1 nodes contribute $O(\Delta^{1/2})$ to $\|\Psi_\Delta-\hat\Psi\|_\Delta$. The same remark holds for ψ_Δ below, and for Υ_Δ at the last node of each line.) The same argument, with the second part of Lemma 5.4, gives $\|\Upsilon_\Delta-\hat\Upsilon\|_\Delta\to 0$. Next, by (5.4), the transpose of $f\mapsto y$ has norm at most $L^2/4$. So it is enough to compare the transpose applied to $\chi_\Theta\hat\Upsilon$ with $\hat\psi$. The function $\hat\psi$ vanishes on $\{x=0\}$, $\{x=L\}$ and at the end of the lines, and Lemma 5.4 gives $(f,\hat\psi)_\Delta=(y,\mathsf{S}^*\hat\psi)_\Delta$. Moreover, $\mathsf{S}^*\hat\psi-\chi_\Theta\hat\Upsilon=-\zeta'_\psi+\varrho_\psi$ at the nodes with $n<N_t$, $j<N_a$ and $x_i\notin\partial\Theta$. This difference is bounded at the nodes $x_i\in\partial\Theta$, and it is equal to $-\chi_\Theta\hat\Upsilon$ at the last node of each line. There are fewer than N_t+N_a last nodes, so they contribute $O(\Delta^{1/2})$ to the norm $\|\cdot\|_\Delta$. The nodes $x_i\in\partial\Theta$ contribute $O(\delta x^{1/2})$. Hence $|(f,\psi_\Delta-\hat\psi)_\Delta|\le\frac{L^2}{4}\big(\|\Upsilon_\Delta-\hat\Upsilon\|_\Delta+\|\mathsf{S}^*\hat\psi-\chi_\Theta\hat\Upsilon\|_\Delta\big)\|f\|_\Delta+o(1)\|f\|_\Delta$, and $\|\psi_\Delta-\hat\psi\|_\Delta\to 0$. Finally, $\Psi+\psi$ is continuous, so $\mathsf{P}_\Delta\big(\chi_\omega(\hat\Psi+\hat\psi)\big)\to(\Psi+\psi)|_{U\times\omega}$ in X, and $\mathsf{Q}_\Delta g-\hat g\to 0$ uniformly. This proves (A3). □

With Proposition 5.5, the convergence results below (Proposition 5.6 and Theorem 5.7) only depend on assumption (C), which holds in our experiments. Theorem 5.7 also needs the nondegeneracy of the continuous problem at h_ε. The proof of Proposition 5.5 is qualitative. Rates can be obtained in the same way when the derivatives are Lipschitz continuous. But χ_Θ is discontinuous, and this gives

residuals of order one at the nodes next to $\partial\Theta$. So this argument only gives the rate $O(\delta x^{1/2})$. This rate is not sharp, since the observed order is one (Table 1).

5.5. **Convergence of the discrete controls.**

Proposition 5.6. *Assume* (C), *and let* $\varepsilon > 0$. *Then* $\mathsf{P}_\Delta h_\varepsilon^\Delta \to h_\varepsilon$ *in* X, *the discrete traces converge,* $\mathsf{E}_\Delta(\mathcal{L}_\Delta h_\varepsilon^\Delta + b_\Delta) \to \mathcal{L}h_\varepsilon + b$ *in* H, *and the discrete sensitivities* $\Phi^\Delta_{h_\varepsilon^\Delta}(e)$ *converge to* $\Phi_{h_\varepsilon}(e)$, $e \in \{0, L\}$.

Proof. We have $\frac{1}{2}\|h_\varepsilon^\Delta\|^2 \le K_\varepsilon^\Delta(h_\varepsilon^\Delta) \le K_\varepsilon^\Delta(0) = \frac{1}{2\varepsilon}\|b_\Delta\|^2$. So the functions $u_\Delta = \mathsf{P}_\Delta h_\varepsilon^\Delta$ are bounded in X. Let u be a weak limit point. For $g \in H$,

$$(\mathsf{E}_\Delta \mathcal{L}_\Delta h_\varepsilon^\Delta, g)_H = (h_\varepsilon^\Delta, \mathcal{L}_\Delta^* \mathsf{Q}_\Delta g)_{X_\Delta} = (u_\Delta, \mathsf{P}_\Delta \mathcal{L}_\Delta^* \mathsf{Q}_\Delta g)_X \to (u, \mathcal{L}^* g)_X = (\mathcal{L}u, g)_H,$$

so $\mathsf{E}_\Delta \mathcal{L}_\Delta h_\varepsilon^\Delta \rightharpoonup \mathcal{L}u$ in H along the corresponding subsequence. By weak lower semicontinuity of the norms, $K_\varepsilon(u) \le \liminf K_\varepsilon^\Delta(h_\varepsilon^\Delta)$. On the other hand, by (A2), $K_\varepsilon^\Delta(h_\varepsilon^\Delta) \le K_\varepsilon^\Delta(\mathsf{R}_\Delta h_\varepsilon) \to K_\varepsilon(h_\varepsilon)$. Hence $K_\varepsilon(u) \le K_\varepsilon(h_\varepsilon)$, and $u = h_\varepsilon$ because the minimizer is unique. So the whole family converges weakly, and $K_\varepsilon^\Delta(h_\varepsilon^\Delta) \to K_\varepsilon(h_\varepsilon)$. Each of the two terms of K_ε^Δ is weakly lower semicontinuous along the family, and their sum converges. So each term converges to the corresponding term of $K_\varepsilon(h_\varepsilon)$. Weak convergence and convergence of the norms give the strong convergences. Finally, E_Δ is an isometry. So $\Phi^\Delta_{h_\varepsilon^\Delta}(e)$ is the $L^2(U)$ inner product of two components of $\mathsf{E}_\Delta(\mathcal{L}_\Delta h_\varepsilon^\Delta + b_\Delta)$, which converge strongly. □

Theorem 5.7. *Assume* (C) *and let* $\varepsilon > 0$. *Assume that* $\nabla\Phi_{h_\varepsilon}(0)$ *and* $\nabla\Phi_{h_\varepsilon}(L)$ *are linearly independent and that* (4.3) *holds at* $h_0 = h_\varepsilon$ *with* $g_k = \nabla F_k(h_\varepsilon)/\|\nabla F_k(h_\varepsilon)\|$. *Let* $h^* = h(c)$ *be the exact control of Proposition 4.2. Then, for* Δ *small, the discrete criterion holds at* h_ε^Δ, *the discrete exact control* $h^{\Delta,*}$ *is well defined, and* $\mathsf{P}_\Delta h^{\Delta,*} \to h^*$.

Proof. By Proposition 5.6, the discrete gradients $\mathcal{L}_\Delta^* w_k^\Delta$ converge to $\nabla F_k(h_\varepsilon) \neq 0$. So the discrete directions converge, and the traces of the associated solutions also converge. The entries of M, S_k and $G(0)$ are inner products of these traces. Hence the discrete quantities converge, $\Gamma^\Delta \to \Gamma < 1$, and Proposition 4.2 applies for Δ small. We choose $r > \rho$ with $\|M^{-1}\|(|G(0)| + qr^2) < r$ and $2qr\|M^{-1}\| < 1$. For Δ small, the discrete fixed point maps $\mathcal{T}^\Delta$ are $\bar\Gamma$-contractions of $\overline{B}(0, r)$ with $\bar\Gamma < 1$, and $|c^\Delta - c| \le \bar\Gamma|c^\Delta - c| + |\mathcal{T}^\Delta(c) - \mathcal{T}(c)|$. The last term tends to zero. Hence $c^\Delta \to c$ and $\mathsf{P}_\Delta h^{\Delta,*} \to h^*$. □

6. Numerical study in dimension one

6.1. **Data.** We take $L = T = A = 1$, $\mu(a) = 0.2 + a$, $\kappa = 1$, the seasonal arrival of individuals of age close to 0.25,

$$\xi(t, a, x) = 10\, e^{-(a-0.25)^2/0.02}\Big(1 + 0.5\sin(2\pi t)\Big),$$

$\Theta = (0.45, 0.65)$ and $\omega = (0.10, 0.25) \cup (0.75, 0.88)$. The set ω meets the two components of $\Omega \setminus \overline{\Theta}$ (Figure 1). The deformation modes are $V_0(x) = -\vartheta((0.1 - x)/0.1)$ and $V_L(x) = \vartheta((x - 0.9)/0.1)$, where ϑ is a smooth step from 0 to 1. The mode V_0 moves only the end point 0, and V_L moves only the end point L. We use $N_t = N_a = N_x = N$, with $N = 80$ if nothing else is said. We stop

the conjugate gradient method when the relative residual is 10^{-9}, and Newton's method when $|G(c)|_\infty < 10^{-14}$. All the computations were done in Python, with NumPy and SciPy, on one core. For $N = 80$, one exact control takes about 36 seconds (Table 5).

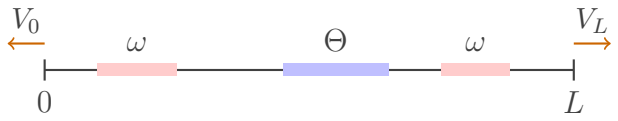


FIGURE 1. Geometry in dimension one. Habitat $\Omega = (0, L)$, control region ω (red), observation region Θ (blue) and deformation modes.

6.2. **The Hadamard formula.** We compare the discrete sensitivity $\Phi_0^\Delta(e)$ with the derivative D_e^Δ of $\tau \mapsto J_0^\Delta(\Omega_\tau)$. This derivative is computed on stretched grids by the complex-step method [24], which is exact up to rounding errors (Table 1). The difference $\mathcal{E}_\Delta(e) = D_e^\Delta - \Phi_0^\Delta(e)$ is divided by about 2 when N is doubled. So the discrete Hadamard formula is consistent at first order, which agrees with the first-order accuracy of (5.1). A Richardson extrapolation gives $\Phi_0(0) \approx 0.0500$ and $\Phi_0(L) \approx 0.0644$. Since $\xi \geq 0$, both values are positive, as Proposition 3.1 predicts. This means that a larger habitat gives a larger observation.

TABLE 1. Discrete Hadamard formula in dimension one ($h = 0$). Sensitivity $\Phi_0^\Delta(e)$, derivative D_e^Δ of the discrete observation (complex step), and difference $\mathcal{E}_\Delta(e)$.

N	$\Phi_0^\Delta(0)$	D_0^Δ	$\mathcal{E}_\Delta(0)$	$\Phi_0^\Delta(L)$	D_L^Δ	$\mathcal{E}_\Delta(L)$
20	0.04220	0.04364	1.44×10^{-3}	0.05358	0.05543	1.84×10^{-3}
40	0.04590	0.04658	6.85×10^{-4}	0.05866	0.05955	8.90×10^{-4}
80	0.04789	0.04822	3.31×10^{-4}	0.06145	0.06188	4.34×10^{-4}
160	0.04893	0.04909	1.62×10^{-4}	0.06291	0.06312	2.14×10^{-4}
320	0.04946	0.04954	8.03×10^{-5}	0.06366	0.06377	1.06×10^{-4}

6.3. **Approximate insensitizing.** Table 2 and Figure 2 show the penalized controls. The sensitivities decrease roughly like ε, as expected from (4.2). They reach 10^{-7} for $\varepsilon = 10^{-5}$. At the same time, the norm of the control grows only from 8.0 to 8.2 between $\varepsilon = 10^{-3}$ and 10^{-5}. The traces of y are reduced by four orders of magnitude, but the traces of z only by a factor four. So the control acts on the flux of individuals through the boundary. The number of conjugate gradient iterations grows only like $\varepsilon^{-0.3}$. The condition number of $I + \varepsilon^{-1}\mathcal{L}_\Delta^*\mathcal{L}_\Delta$ is $1 + \|\mathcal{L}_\Delta\|^2/\varepsilon$, and it would suggest up to $O(\varepsilon^{-1/2})$ iterations. The reason for the smaller number is that the singular values of $\mathcal{L}_\Delta$ form two well separated groups, one for the traces of y and one for the traces of z (Section 6.8). For the same reason, the number of iterations depends only a little on the grid (Table 3), and we do not need a preconditioner. The discrete map $\mathcal{L}_\Delta$ is a matrix, so these results

do not tell us if b belongs to the range of $\mathcal{L}$. If it does, the quantity $\varepsilon^{-1}\|r_\varepsilon\|^2$ tends to zero. This quantity has not started to decrease at $\varepsilon = 10^{-5}$, so we are not yet in the asymptotic regime of Proposition 4.1. Finally, for the penalized control, the discrete sensitivity agrees with the derivative of the discrete observation on the deformed habitats. For $\varepsilon = 10^{-3}$ and $N = 80$, the derivatives D_0^Δ and D_L^Δ are 1.08×10^{-4} and 1.66×10^{-4}, and $\Phi_{h_\varepsilon}^\Delta(0) = 1.07 \times 10^{-4}$ and $\Phi_{h_\varepsilon}^\Delta(L) = 1.63 \times 10^{-4}$.

TABLE 2. Penalized controls ($N = 80$). Conjugate gradient iterations, sensitivities, norms of the traces and of the control.

ε	CG	$\lvert\Phi(0)\rvert$	$\lvert\Phi(L)\rvert$	$\lVert\partial_n y\rVert$	$\lVert\partial_n z\rVert$	$\lVert h_\varepsilon\rVert$
$h = 0$	–	4.8×10^{-2}	6.1×10^{-2}	2.54	4.8×10^{-2}	0
10^{-1}	12	1.4×10^{-2}	1.9×10^{-2}	1.26	3.0×10^{-2}	3.99
10^{-2}	30	1.2×10^{-3}	1.8×10^{-3}	0.24	1.7×10^{-2}	7.28
10^{-3}	61	1.1×10^{-4}	1.6×10^{-4}	2.7×10^{-2}	1.4×10^{-2}	8.00
10^{-4}	97	9.6×10^{-6}	1.5×10^{-5}	2.6×10^{-3}	1.4×10^{-2}	8.08
10^{-5}	168	2.5×10^{-7}	2.1×10^{-7}	1.7×10^{-4}	1.2×10^{-2}	8.21

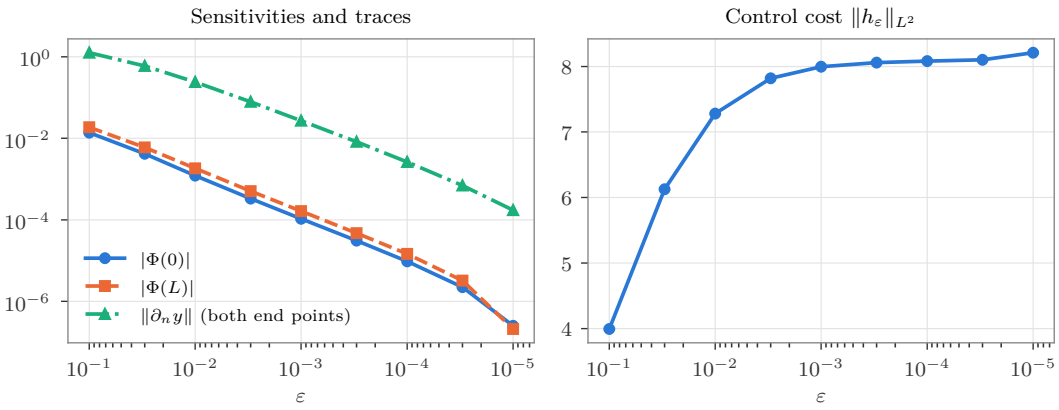


FIGURE 2. Penalized controls h_ε ($N = 80$) as functions of ε, in logarithmic scales. Left, the sensitivities $|\Phi_{h_\varepsilon}(0)|$, $|\Phi_{h_\varepsilon}(L)|$ and the norm of the traces of the controlled state. Right, the norm of the control.

6.4. **Exact insensitizing.** Now we apply the two-stage algorithm of Section 4.2 with the two end-point modes ($K = 2$). We start from h_ε with $\varepsilon = 10^{-3}$. Table 3 shows that, for all the grids, Newton's method converges quadratically in three iterations. It cancels the discrete sensitivities up to rounding errors. The correction changes the norm of the control by less than one percent.

The discrete sensitivity is not exactly the derivative of the discrete observation (Table 1). We want to check that the computed controls also insensitize the discrete problem posed on the deformed habitats. For this, we computed with the complex-step method the derivatives D_0^Δ and D_L^Δ of $\tau \mapsto J_h^\Delta(\Omega_\tau)$ for

the exact control h^*. Since $\Phi^\Delta_{h^*} = 0$, D^Δ_e is only the consistency error of the discrete Hadamard formula. This error depends quadratically on the solutions, and the control has reduced their traces. Therefore the derivatives are smaller than without control, by a factor between 6.6×10^3 and 1.4×10^5. They tend to zero when the grid is refined, and the observed orders are between 1.15 and 1.53. The computed controls also converge. The relative L^2 difference between $\int_\omega h^*\,dx$ computed with $N = 20$, 40, 80 and with $N = 160$ is 7.5%, 4.4% and 2.6%. The observed order is about 0.77. This order must be taken with care, because the reference solution is also approximate.

TABLE 3. Exact insensitizing controls h^* (two-stage algorithm, $\varepsilon = 10^{-3}$). Iterations, sensitivities before and after the correction, criterion (4.3), norm of the control, and derivatives of the discrete observation on deformed habitats with $h = 0$ and $h = h^*$.

N	CG	Newton	$\lvert\Phi_{h_\varepsilon}\rvert_\infty$	$\lvert\Phi_{h^*}\rvert_\infty$	$\lVert h^*\rVert$	$\lvert D\rvert_\infty$, $h = 0$	$\lvert D\rvert_\infty$, h^*
20	37	3	1.3×10^{-4}	9.6×10^{-16}	8.08	5.5×10^{-2}	8.5×10^{-6}
40	48	3	1.4×10^{-4}	1.5×10^{-16}	7.92	6.0×10^{-2}	3.8×10^{-6}
80	61	3	1.6×10^{-4}	1.9×10^{-16}	8.05	6.2×10^{-2}	1.5×10^{-6}
160	80	3	1.8×10^{-4}	2.2×10^{-16}	8.12	6.3×10^{-2}	6.0×10^{-7}

We cannot skip the first stage. If we start from $h_0 = 0$, Newton's method does not converge. For $N = 80$, the sensitivities oscillate between 2×10^{-3} and 6, and there is no convergence after fifty iterations. But the first stage does not need to be accurate. Table 4 gives the quantities of Proposition 4.2. At $h_0 = 0$, we have $\Gamma \approx 30$. If we start from h_ε with $\varepsilon = 10^{-1}$, 10^{-2} and 10^{-3}, Newton's method converges in 7, 4 and 3 iterations, and the norm of the exact control is 7.81, 7.82 and 8.05 ($N = 80$). The criterion (4.3) holds only for $\varepsilon = 10^{-3}$. So this criterion is sufficient but not necessary. The matrix M and the constant q change little with ε. On the contrary, $|G(0)|$, and so Γ, is roughly proportional to ε. All these quantities are stable when the grid changes. The values $\Gamma^\Delta = 0.305$ and 0.297 for $N = 40$ and 80 support the assumption of Theorem 5.7. The angle θ between the two gradients satisfies $\cos\theta = 0.30$ for $\varepsilon = 10^{-3}$, and 0.68–0.69 at $h_0 = 0$. The condition number of M is between 1.9 and 5.6. So the gradients are far from collinear. The large values of $\|M^{-1}\|$ come from the small norms of the gradients, about 3×10^{-3} for $\varepsilon = 10^{-3}$.

6.5. **Qualitative behaviour of the controlled system.** Figure 3 shows the effect of the controls for finite deformations. Without control, $J_h(\Omega_\tau)$ varies linearly, by about 3% when an end point moves by 0.01. With the exact control, the variation is quadratic, and it is smaller than 1% for $|\tau| \le 0.02$. The penalized control with $\varepsilon = 10^{-3}$ gives almost the same curve. The controls also reduce the observation itself ($J_{h^*} = 1.65 \times 10^{-3}$), even if this was not asked. The exact control (Figure 5) is concentrated on the ages of the arriving individuals, and it is negative near the uncertain boundary. It acts against the recruitment, and this reduces the flux of individuals through the boundary (Figure 4). The maximum

TABLE 4. Quantities of Proposition 4.2 at the starting control h_0 of the correction, and number of Newton iterations ("–" means no convergence in 30 iterations). The radius ρ is a bound for the correction only when $\Gamma < 1$.

N	h_0	$\lvert G(0)\rvert$	$\lVert M^{-1}\rVert$	q	Γ	ρ	Newton
40	0	7.4×10^{-2}	237	2.0×10^{-3}	33.3	35.3	–
40	h_ε, $\varepsilon = 10^{-1}$	2.2×10^{-2}	347	1.9×10^{-3}	19.5	15.1	7
40	h_ε, $\varepsilon = 10^{-2}$	2.0×10^{-3}	511	1.5×10^{-3}	3.24	2.08	4
40	h_ε, $\varepsilon = 10^{-3}$	1.8×10^{-4}	559	1.3×10^{-3}	0.305	0.203	3
80	0	7.8×10^{-2}	226	2.0×10^{-3}	31.2	35.1	–
80	h_ε, $\varepsilon = 10^{-1}$	2.3×10^{-2}	327	1.8×10^{-3}	18.3	15.2	7
80	h_ε, $\varepsilon = 10^{-2}$	2.2×10^{-3}	485	1.5×10^{-3}	3.13	2.13	4
80	h_ε, $\varepsilon = 10^{-3}$	2.0×10^{-4}	534	1.3×10^{-3}	0.297	0.208	3

of $|\partial_n y(\cdot, \cdot, L)|$ goes down from 5.7 to 0.14. By Proposition 3.1, the controlled source must take negative values, since the exact control cancels the sensitivities. Here the controlled state also changes sign. On the whole domain, its values are between -0.14 and 0.47, while the uncontrolled state is nonnegative with maximum 1.36. Figure 6 shows the state at $t = T/2$. So this control makes sense when y is the deviation from a reference population larger than 0.14. We also imposed the constraint $\xi + h\chi_\omega \geq 0$, and minimized (4.1) by L-BFGS-B with first-order normal derivatives. In this case the optimal control is $h = -\xi$ in ω, that is, it suppresses all the arrivals in ω. The sensitivities are only reduced by 44–46%, from $(4.71, 6.05) \times 10^{-2}$ to $(2.55, 3.39) \times 10^{-2}$. They stay positive, as Proposition 3.1 says.

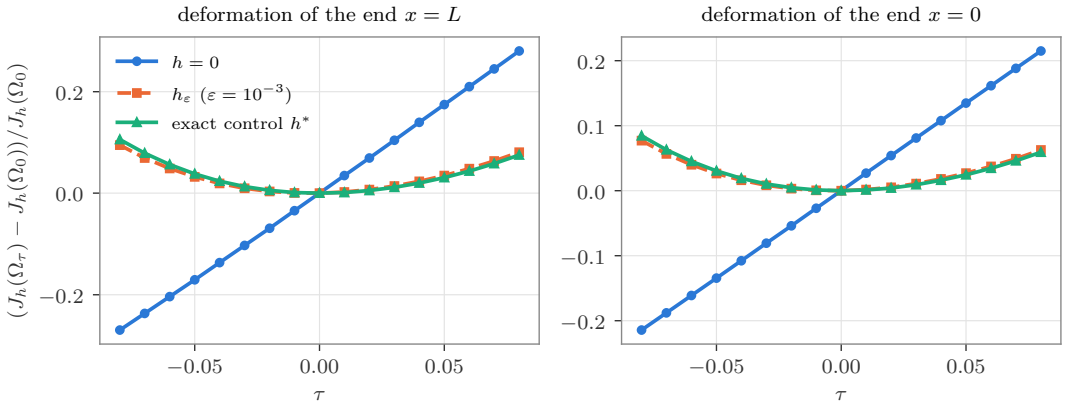


FIGURE 3. Relative variation of $J_h(\Omega_\tau)$ when each end point is deformed. Without control, with the penalized control ($\varepsilon = 10^{-3}$) and with the exact control ($N = 80$).

6.6. **The exact control of minimal norm and comparison of methods.** The exact controls of Table 3 are not unique. If we start the correction from

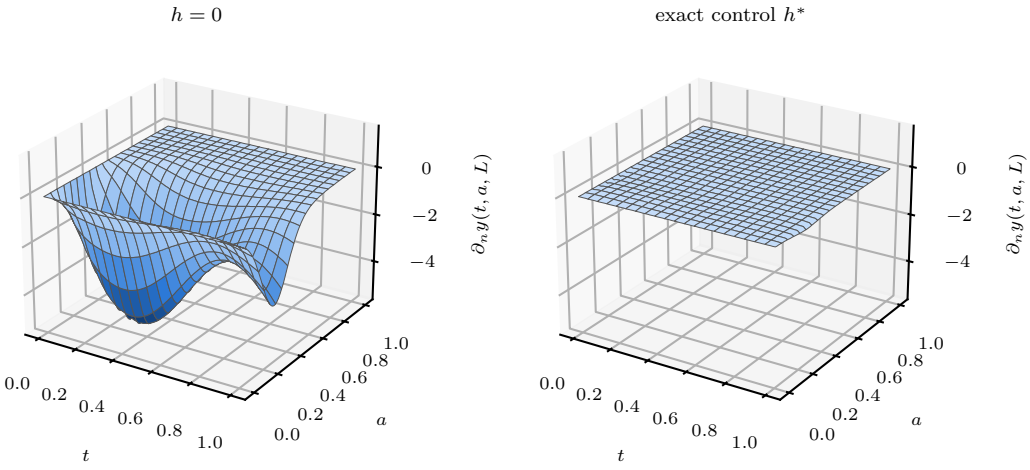


FIGURE 4. Flux of individuals through $x = L$: outward normal derivative $\partial_n y(t, a, L)$, which is negative for an outgoing flux, without control and with the exact control ($N = 80$).

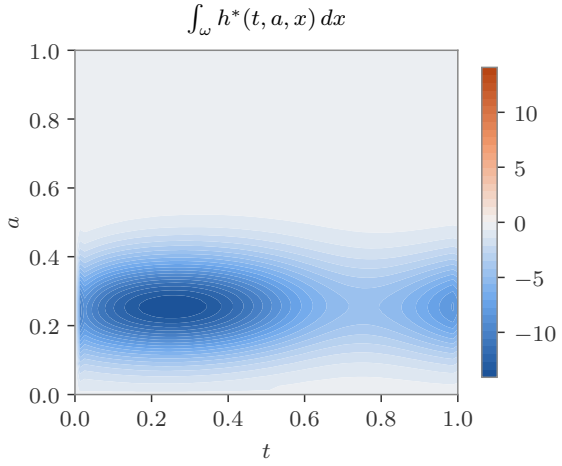


FIGURE 5. The exact insensitizing control h^*, integrated over ω, as a function of time and age ($N = 80$).

h_ε with $\varepsilon = 10^{-2}$ and with 10^{-3}, we obtain two exact controls. Their norms differ by 3%, and their relative distance is 7%. For this reason we computed the exact control of minimal norm $h^\dagger$ (Section 4.3), and we compared several ways to obtain it (Table 5). We measure the cost by the number of forward or backward recursions along the characteristic lines. Each recursion costs the same as one solution of the state equation.

We start the Lagrange–Newton method of Section 4.3 from the two-stage control with $\varepsilon = 10^{-2}$. It converges in seven updates of the multipliers to a control $h^\dagger$ of norm 7.42. The residual of the Lagrange conditions (4.5) is of order 10^{-11}. For $N = 40$, we started it from the three two-stage controls obtained with $\varepsilon = 10^{-1}$,

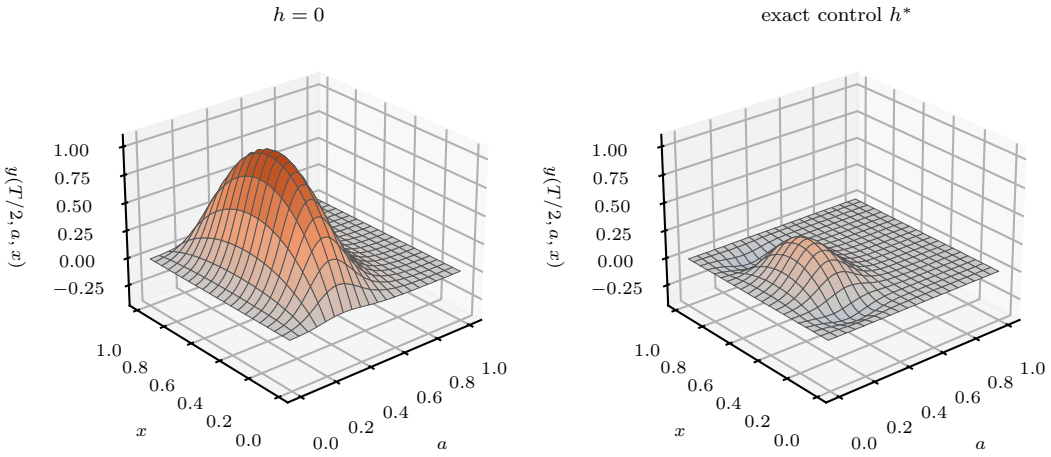


FIGURE 6. The state at $t = T/2$, $y(T/2, a, x)$, without control and with the exact control ($N = 80$). At $t = T/2$, the controlled state is between -0.09 and 0.35 (maximum 1.10 without control). Its negative values are in and near the two components of ω, and the minimum is at $x \approx 0.84$.

10^{-2} and 10^{-3}, and it converges to the same control, up to a relative difference of 2×10^{-11}. The controls given by the two-stage algorithm are at a relative distance of 26–33% from $h^\dagger$, but their norms are only 5–8% larger. We also minimized directly $\frac{1}{2}\sum_k F_k(h)^2 + \frac{\eta}{2}\|h\|^2$ by L-BFGS, starting from $h = 0$. The result approaches $h^\dagger$ when $\eta \to 0$, but the sensitivities are only reduced to 10^{-3}–10^{-7} (Table 5). Then we applied the exact correction to the result for $\eta = 10^{-8}$. This gives an exact control within 0.08% of $h^\dagger$, for less than a quarter of the cost of the Lagrange–Newton method. Note that the count of the Lagrange–Newton method in Table 5 does not include the computation of its starting control. In our experiments, this last combination is the most efficient way to compute $h^\dagger$. The two-stage algorithm stays the cheapest way to obtain one exact control.

To check the assumption of Proposition 4.3, we computed the two smallest eigenvalues of the discrete self-adjoint operator $\sum_k \lambda_k \mathcal{A}_k$ at $h^\dagger$. We used the implicitly restarted Lanczos method, and each product costs four recursions along the lines. For $N = 20$ and $N = 40$, the smallest eigenvalue of $I + \sum_k \lambda_k \mathcal{A}_k$ is 0.886 and 0.892. We did not do this computation for $N = 80$, because it did not finish in seventy minutes. So, for $N = 20$ and $N = 40$, and up to the accuracy of the eigenvalue computation, $h^\dagger$ is the unique solution of the discrete problem (4.4). Moreover every discrete exact control h satisfies $\|h\|^2 \geq \|h^\dagger\|^2 + 0.88\,\|h - h^\dagger\|^2$. This does not prove the assumption of Proposition 4.3 for the continuous problem, or for finer grids. The fact that the eigenvalue is stable when the grid changes is only an indication in this direction. Finally, $h^\dagger$ is stable with respect to the grid. Its norm is 7.47, 7.30 and 7.42 for $N = 20$, 40 and 80. The relative difference between $\int_\omega h^\dagger\,dx$ computed with $N = 20$, 40 and with $N = 80$ is 3.7% and 1.3%.

TABLE 5. Computation of exact insensitizing controls ($N = 80$). The column "recurs." gives the number of forward or backward recursions along all the characteristic lines. Each recursion costs the same as one solution of the state equation (5.1). The other columns give the CPU time, the sensitivity, the norm of the control and the relative distance $\|h - h^\dagger\|/\|h^\dagger\|$ to the exact control of minimal norm $h^\dagger$.

method	recurs.	CPU (s)	$\lvert\Phi\rvert_\infty$	$\|h\|$	dist. to $h^\dagger$
two-stage, $\varepsilon = 10^{-2}$	138	17	4.5×10^{-15}	7.821	2.6×10^{-1}
two-stage, $\varepsilon = 10^{-3}$	262	36	1.9×10^{-16}	8.047	3.3×10^{-1}
Lagrange–Newton	2174	237	6.8×10^{-14}	7.421	0
L-BFGS, $\eta = 10^{-6}$	126	14	1.3×10^{-3}	6.913	8.1×10^{-2}
L-BFGS, $\eta = 10^{-8}$	474	54	1.7×10^{-5}	7.413	1.3×10^{-3}
L-BFGS, $\eta = 10^{-10}$	754	85	1.7×10^{-7}	7.420	2.0×10^{-4}
L-BFGS, $\eta = 10^{-8}$, + corr.	484	54	1.2×10^{-18}	7.421	7.8×10^{-4}

6.7. **Influence of the geometry.** When $\omega \cap \Theta = \emptyset$, result (R1) asks that ω meets the two components of $\Omega \setminus \overline{\Theta}$ that touch the boundary. To see what happens in the other case, we took $\omega = (0.10, 0.25)$, which only meets the component that touches 0 ($N = 60$). Figure 7 compares this case with the previous one. When ε goes from 10^{-1} to 10^{-5}, the trace of y at the end point 0, on the side of ω, does not decrease monotonically, and it is only divided by 86. The trace at the other end point L is only divided by 13. When ω meets both components, both traces are divided by about 7000. At the same time, the norm of the control grows from 3.2 to 97, instead of staying close to 8, and the number of conjugate gradient iterations grows from 11 to 519. The control can act on the far boundary only through the observation region, by diffusion. Section 6.8 shows that, at the discrete level, the traces at the far end point and the traces of the adjoint state cannot be controlled independently. Result (R1) does not cover this configuration. As far as we know, it is an open question whether approximate insensitizing holds in this case. The experiment gives a partial answer. If it holds, it is very costly, and the far end point is the main difficulty. In practice, this means that the control region should be placed on both sides of the monitored region.

6.8. **Spectrum of the control-to-traces map.** To explain these observations, we computed all the singular values $s_1 \geq s_2 \geq \dots$ of $\mathcal{L}_\Delta$, for the norms of Section 5, so that $\|\mathcal{L}_\Delta\| = s_1$. We also computed the coefficients $|u_k^\top b_\Delta|$ of the uncontrolled traces on the corresponding left singular vectors u_k. We did this for $N = 20$ and $N = 40$, and for the two control regions of Section 6.7. We built $\mathcal{L}_\Delta \mathcal{L}_\Delta^*$ column by column, with one backward and one forward recursion for each column, and we computed its eigenvalues. With this method, singular values smaller than about 10^{-8} are not resolved. The trace space has dimension $4(N+1)^2$. The components of the singular vectors show which traces they involve. Figure 8 shows the results.

Two-sided control region. For $N = 40$, the $2N^2$ largest singular values are in $[0.082, 0.352]$, and for $N = 20$ they are in $[0.112, 0.345]$. The corresponding

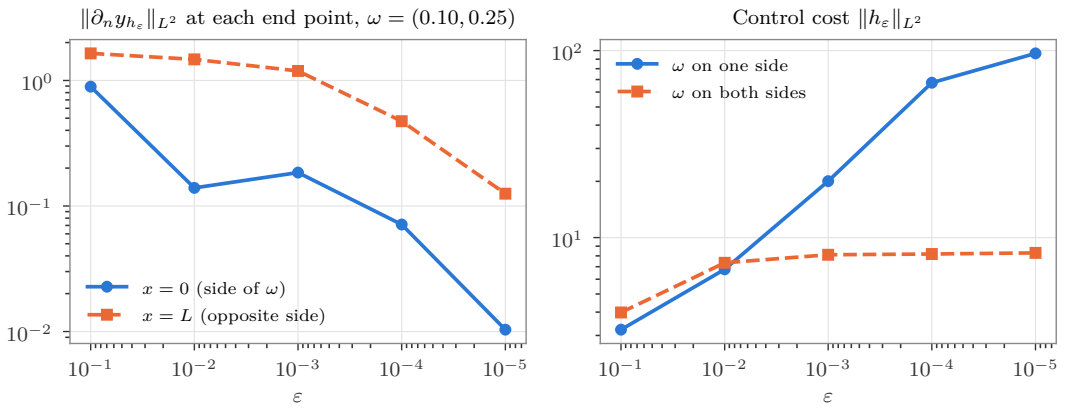


FIGURE 7. Control region on one side of Θ, $\omega = (0.10, 0.25)$ ($N = 60$). Left, the norm of the trace of y at each end point. Right, the norm of the control, compared with the two-sided region $\omega = (0.10, 0.25) \cup (0.75, 0.88)$.

singular vectors only involve the traces of y. The next singular values only involve the traces of z, and they are smaller than 1.4×10^{-3}. So the traces of y are well controlled, and the traces of z only weakly, as observed in Section 6.3. This structure explains the behaviour of the conjugate gradient method. For $N = 40$ and $\varepsilon \geq 10^{-5}$, the eigenvalues of $I + \varepsilon^{-1}\mathcal{L}_\Delta^*\mathcal{L}_\Delta$ are of two kinds. There is a group of $2N^2$ eigenvalues in $[1 + s_{2N^2}^2/\varepsilon, 1 + s_1^2/\varepsilon]$, and the ratio of these eigenvalues is bounded by $(s_1/s_{2N^2})^2 \approx 18$, independently of ε. The other eigenvalues are in $[1, 1.2]$. The number of iterations depends on this ratio, and not on the condition number. The lower end of the first group decreases slowly when the grid is refined. This agrees with the slow growth of the number of iterations with N. The number of singular values larger than 0.2 grows from 196 to 511 when N goes from 20 to 40. This agrees with the fact that $\mathcal{L}$ is not compact (Remark 4.4). Finally, the coefficients $|u_k^\top b_\Delta|$ decay with the singular values. For $N = 40$, the ratios $|u_k^\top b_\Delta|/s_k$, $s_k \geq 10^{-8}$, are smaller than 1.9, and their median is 2.8×10^{-3}. The component of b_Δ on the singular vectors with $s_k < 10^{-8}$ has relative norm 1.4×10^{-7}. So the discrete Picard condition [15] holds. This agrees with the assumption that b belongs to the range of $\mathcal{L}$, but it does not prove it.

One-sided control region. For $N = 40$, the singular values form three groups. There are N^2 values in $[0.083, 0.301]$, for the traces of y at the end point 0, on the side of ω. There are N^2 values in $[1.8 \times 10^{-4}, 1.7 \times 10^{-2}]$, for the traces of y at the far end point L. The other values are below the resolution of the computation, and they correspond to the traces of z. So the numerical rank of $\mathcal{L}_\Delta$ is $2N^2$, which is half the number $4N^2$ of nontrivial traces. The same holds for $N = 20$. The component of b_Δ orthogonal to the range of $\mathcal{L}_\Delta$ has relative norm 1.9×10^{-2}. Therefore, for every control, the residual $\|\mathcal{L}_\Delta h + b_\Delta\|$ is at least 1.9% of $\|b_\Delta\|$. The projection of b_Δ on the singular vectors with $s_k < 0.05$ has norm $0.59\|b_\Delta\|$. The control can reduce this part only at a high cost, since the ratios $|u_k^\top b_\Delta|/s_k$ reach 14. This explains the results of Section 6.7.

Remark 6.1. In the one-sided case, the rank deficiency has a simple structural reason. Let $\omega = (\alpha, \beta)$, and let x_p and x_r be the first and the last nodes in ω. Assume that the nodes used in the formulas for the normal derivatives are outside $\overline{\omega}$. Here $x_2 < \alpha$ as soon as $N \geq 20$. Recall that $\mathcal{L}_\Delta$ is defined with $\xi = 0$. At the nodes x_i with $i < p$, and at the nodes with $i > r$, the source of (5.1) is zero. So, by induction along each characteristic line, the value of y at these nodes is a linear function of the value of y at x_p, respectively at x_r, on the same line. Since $\overline{\Theta} \subset (\beta, L)$, the traces of y at L and the source $\chi_\Theta y$ of the backward recursion in (5.1) are linear functions of the N^2 values $(y_j^n)_r$, $1 \leq n, j \leq N$. The same is true for the traces of z at both end points. The traces of y at 0 are linear functions of the values $(y_j^n)_p$. Therefore the rank of $\mathcal{L}_\Delta$ is at most $2N^2$. In the continuous setting, the same argument shows that the traces of y_h at L and of z_h at both end points only depend on $y_h(\cdot, \cdot, \beta)$. This suggests that the range of $\mathcal{L}$ is not dense in this configuration, but we do not prove it. This would agree with the geometric assumption of (R1). It would not exclude approximate insensitizing, which only asks that the integrals of the products $\partial_n y\, \partial_n z$ are small. But the residual r_ε of Proposition 4.1 would not tend to zero, and the bound (4.2) of Proposition 4.1 would not guarantee small sensitivities. When ω meets D_- and D_+, the same argument uses the values of y at the four end points of the two components of ω, and it gives no such bound.

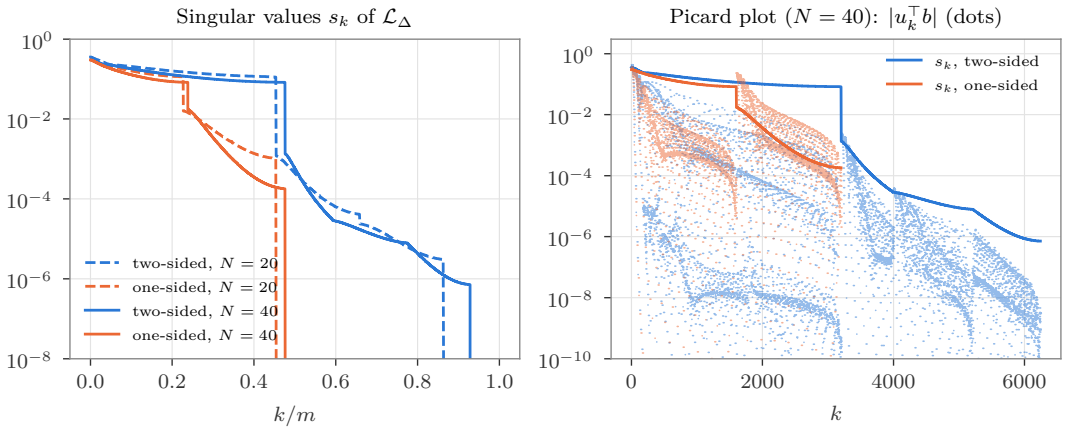


FIGURE 8. Left, the singular values s_k of $\mathcal{L}_\Delta$ as functions of k/m, where $m = 4(N+1)^2$ is the dimension of the trace space. We show the two-sided region $\omega = (0.10, 0.25) \cup (0.75, 0.88)$ and the one-sided region $\omega = (0.10, 0.25)$, for $N = 20$ and $N = 40$. Values below 10^{-8} are not resolved and are not shown. Right, the Picard plot for $N = 40$, with the coefficients $|u_k^\top b_\Delta|$ of the uncontrolled traces (dots) and the singular values (lines).

6.9. **Quantitative study.** Table 6 shows the influence of the parameters. We take $N = 40$, $\varepsilon = 10^{-3}$, and $\tau = 0.04$ for the variation of J. We change one

parameter at a time. The mortality is multiplied by 0.5, 2 or 4. The diffusion coefficient is $\kappa \in \{0.25, 0.5, 2\}$. The observation region is shifted. The control region is made narrower, $(0.15, 0.22) \cup (0.78, 0.85)$, or wider, $(0.05, 0.35) \cup (0.70, 0.95)$.

TABLE 6. Quantitative study ($N = 40$, $\varepsilon = 10^{-3}$). Uncontrolled sensitivities and observation, relative variation δJ_0 of the uncontrolled observation for $\tau = 0.04$ (end point L), conjugate gradient and Newton iterations, norm of the exact control, and variation $\delta J^* = (J_{h^*}(\Omega_\tau) - J_{h^*}(\Omega))/J_0(\Omega)$.

case	$\Phi_0(0)$	$\Phi_0(L)$	J_0	δJ_0 (%)	CG	Newton	$\|h^*\|$	δJ^* (%)
reference	0.0459	0.0587	0.0173	13.9	48	4	7.92	0.19
$\mu \times 0.5$	0.0489	0.0623	0.0179	14.2	49	4	7.93	0.19
$\mu \times 2$	0.0406	0.0523	0.0160	13.3	47	4	7.89	0.19
$\mu \times 4$	0.0323	0.0423	0.0139	12.3	44	4	7.85	0.20
$\kappa = 0.25$	0.5250	0.7672	0.0911	8.2	169	4	7.24	0.19
$\kappa = 0.5$	0.1857	0.2491	0.0442	11.3	79	4	7.61	0.19
$\kappa = 2$	0.0087	0.0108	0.0056	16.0	27	5	8.07	0.18
$\Theta = (0.25, 0.45)$	0.0674	0.0319	0.0148	8.7	48	4	7.92	0.10
$\Theta = (0.55, 0.75)$	0.0319	0.0674	0.0148	18.8	48	4	7.92	0.19
ω narrow	0.0459	0.0587	0.0173	13.9	43	5	11.30	0.15
ω wide	0.0459	0.0587	0.0173	13.9	46	4	5.86	−0.36

These results confirm Section 3 and give it in numbers. The sensitivities decrease when the mortality increases, as Proposition 3.2(i) says, but slowly. If we multiply the mortality by 8, they are reduced by about 34%. The diffusion has a much stronger effect. When κ goes from 0.25 to 2, the sensitivities decrease by a factor 60 to 70 and the observation by a factor 16. The relative variation of the uncontrolled observation increases from 8% to 16%. With a small diffusion, the population leaves the habitat more slowly through the boundary. Then the observation is larger, but relatively less sensitive. The two shifted observation regions are symmetric to each other, and the sensitivities are exchanged, as in Proposition 3.2(iii). The observation is most sensitive to the end point that is closest to Θ. The norm of the exact control and the number of Newton iterations depend very little on the mortality, the diffusion or the observation region. In all cases, $\|h^*\| \approx 7.2$–8.1, and Newton's method converges in 4 or 5 iterations. The remaining variation of the observation for $\tau = 0.04$ is between 0.1% and 0.4% of J_0 in absolute value (about 0.2% in most cases), and it is 8–19% without control. The cost of the control depends mainly on the size of the control region. The norm $\|h^*\|$ is 11.3, 7.9 and 5.9 for the narrow, reference and wide regions, and the norm of the recruitment in ω is 1.7, 2.4 and 3.3. Finally, the number of conjugate gradient iterations grows when the diffusion decreases, from 27 to 169. A possible reason is that the traces then depend more on the details of the control.

7. A TWO-DIMENSIONAL HABITAT

We now take $\Omega = (0, 1)^2$, $\mu(a) = 0.2 + a$ and $\kappa = 1$. The recruitment is $\xi(t, a, x) = 10\, e^{-(a-0.25)^2/0.02}(1 + 0.5 \sin(2\pi t))(1 + 0.5x_1)$. It increases from the

left side to the right side. The observation region is $\Theta = (0.4, 0.6)^2$, and the control region is $\omega = \{x : \ 0.2 < |x - x_c|_\infty < 0.35\}$, which is a square ring around Θ ($x_c = (\frac{1}{2}, \frac{1}{2})$). Each mode V_k translates one side. For example, $V_{\text{right}}(x) = \vartheta((x_1 - 0.9)/0.1)\, e_1$. Then $V_{\text{right}} \cdot n = 1$ on the right side and 0 on the other sides, and $F_{\text{right}}(h)$ is the integral of Φ_h over the right side. The problem has four variables. We use $N_t = N_a = N$ and $N \times N$ interior space nodes, with $N = 20$ and 40 (about 2.8×10^6 unknowns for $N = 40$). The data are symmetric with respect to $x_2 = \frac{1}{2}$, so by Proposition 3.2(iii) the bottom and top sides have the same sensitivity.

Hadamard formula. Without control and for $N = 40$, the side sensitivities are $F = (3.61, 4.49, 4.05, 4.05) \times 10^{-3}$ (left, right, bottom, top). The right side has the largest sensitivity. This is the side where the recruitment is larger. This agrees with the idea of Proposition 3.2(i), but the proposition does not apply directly, because the recruitment and its reflection $x_1 \mapsto 1 - x_1$ are not ordered. We also computed the derivatives of J^Δ for the translations of the right and left sides by centred differences on stretched grids. We obtain 4.58×10^{-3} and 3.68×10^{-3}. For the right side, the differences with F are 4.2, 2.0 and 0.93×10^{-4} for $N = 10$, 20, 40. So the discrete Hadamard formula is consistent at first order, as in dimension one.

Insensitizing. Table 7 gives the penalized controls, which make the whole function Φ_h small. It also gives the exact correction for the four modes, computed from h_ε with $\varepsilon = 10^{-3}$. The results are similar to dimension one. The side sensitivities decrease roughly like ε. The number of conjugate gradient iterations grows moderately and depends very little on the grid. Newton's method in $\mathbb{R}^4$ converges quadratically in four iterations (for $N = 40$, the residuals are 5.1×10^{-6}, 5.6×10^{-7}, 4.5×10^{-9}, 3.0×10^{-13} and 2.4×10^{-21}). On the whole boundary, the pointwise sensitivity is reduced by two orders of magnitude (Figure 9(a)). The exact correction cancels the integrals of this function over the sides, but not the function itself. To cancel the function, we would need infinitely many modes. For the translations of the right and left sides ($N = 40$), the derivatives of J^Δ go from 4.6×10^{-3} and 3.7×10^{-3} down to -3.2×10^{-7} and -2.1×10^{-7}. Under finite translations with $|\tau| \le 0.06$, the variation of the observation becomes quadratic and stays below 0.3% of J_0. Without control, it is 11–12% (Figure 9(b)). As in dimension one, the control is negative near the outer boundary of ω. There, it reduces the flux of arriving individuals towards the boundary. Its size is larger on the right, where the recruitment is larger (Figure 9(c)).

The cost is higher in dimension two. For $N = 40$, one penalized control takes about two minutes, and the exact correction takes eight seconds. Unlike in dimension one, the norm of the control continues to grow slowly when ε decreases (from 8.8 to 13.6 between $\varepsilon = 10^{-2}$ and 10^{-4}). This suggests that the uncontrolled traces are more difficult to reach exactly when the traces are defined on the whole boundary.

Effect of the control on the population. Figure 10 shows the effect of the exact control in the habitat and on the boundary ($N = 40$). Without control, the density at $t = T/2$, integrated over the ages, is positive. Its maximum is 0.256, near the centre of the habitat. Its mean over the right half is 8% larger than its

TABLE 7. Two-dimensional habitat. Number of conjugate gradient iterations, largest side sensitivity $|F|_\infty$, largest pointwise contribution $\max_\sigma |\Phi_h(\sigma)|\,\delta x$, norm of the control and observation. The last rows give the exact correction for the four modes.

N	control	CG	$\lvert F\rvert_\infty$	$\max_\sigma \lvert\Phi_h\rvert\delta x$	$\lVert h\rVert$	J_h
20	$h = 0$	–	3.9×10^{-3}	3.5×10^{-4}	0	2.2×10^{-3}
20	$\varepsilon = 10^{-1}$	10	1.1×10^{-3}	9.9×10^{-5}	4.03	5.3×10^{-4}
20	$\varepsilon = 10^{-2}$	27	3.0×10^{-5}	2.0×10^{-6}	8.81	1.6×10^{-5}
20	$\varepsilon = 10^{-3}$	73	7.2×10^{-6}	1.8×10^{-6}	10.88	4.6×10^{-5}
20	$\varepsilon = 10^{-4}$	197	6.0×10^{-7}	1.0×10^{-6}	13.65	2.6×10^{-5}
40	$h = 0$	–	4.5×10^{-3}	2.1×10^{-4}	0	2.4×10^{-3}
40	$\varepsilon = 10^{-3}$	77	5.1×10^{-6}	7.0×10^{-7}	10.55	2.5×10^{-5}
20	exact, 4 modes	–	7.1×10^{-21}	1.8×10^{-6}	10.96	5.3×10^{-5}
40	exact, 4 modes	–	3.0×10^{-21}	7.2×10^{-7}	10.61	3.0×10^{-5}

mean over the left half, because the recruitment is larger on the right. With the exact control, the density is between -0.092 and 0.004. So the control removes almost all the individuals that arrive in ω. The minimum is reached in ω, and the diffusion spreads the negative values to the rest of the habitat (the minimum outside ω is -0.075). As in dimension one, these negative values are consistent with Proposition 3.1, because the controlled source changes sign. Consider now the outgoing flux through the right side, $-\int_0^A \partial_n y\, da$, as a function of t and x_2. Without control, it follows the seasonal recruitment, and its maximum is 1.65. With the exact control, its maximum is 0.22 and it changes sign. As in dimension one, the control makes the observation insensitive by reducing the flux of individuals through the uncertain boundary.

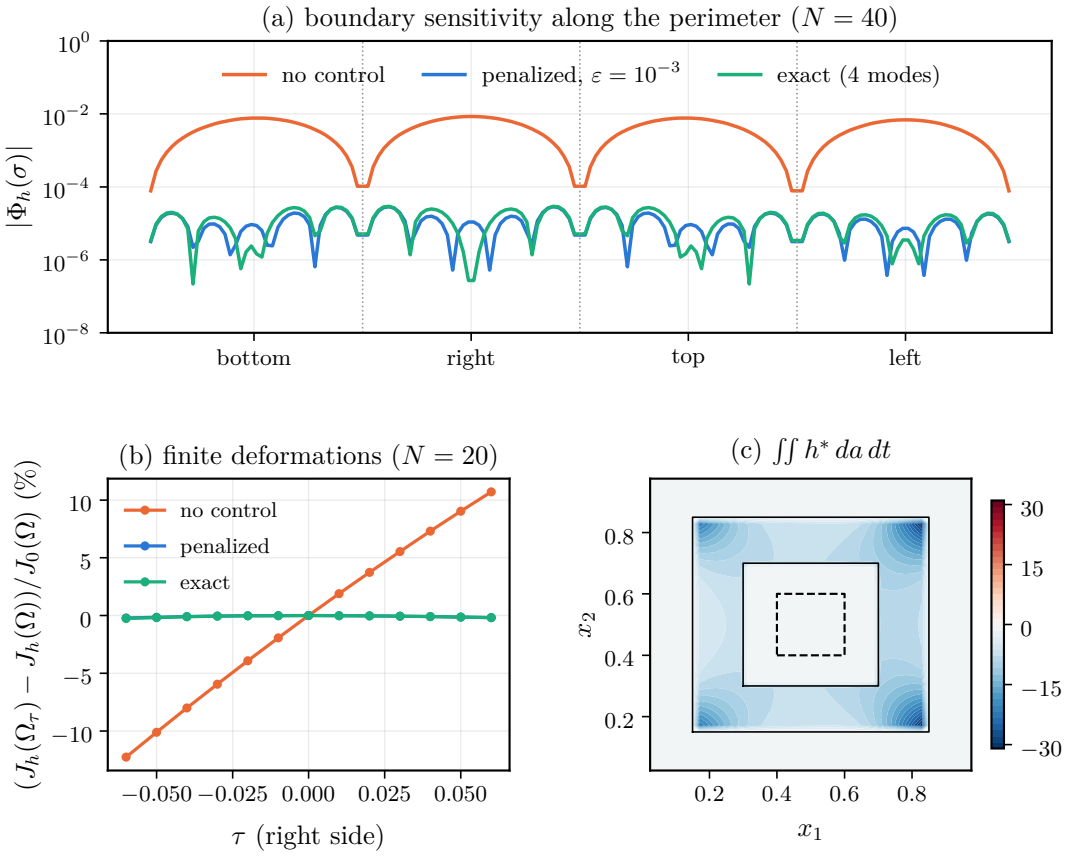


FIGURE 9. Two-dimensional habitat. (a) Pointwise sensitivity $|\Phi_h(\sigma)|$ along the boundary, without control, with the penalized control and with the exact control for the four side modes ($N = 40$). (b) Variation of the observation under a translation of the right side, relative to the observation without control ($N = 20$). (c) Exact control integrated in time and age. The solid lines show the boundary of ω and the dashed line shows Θ.

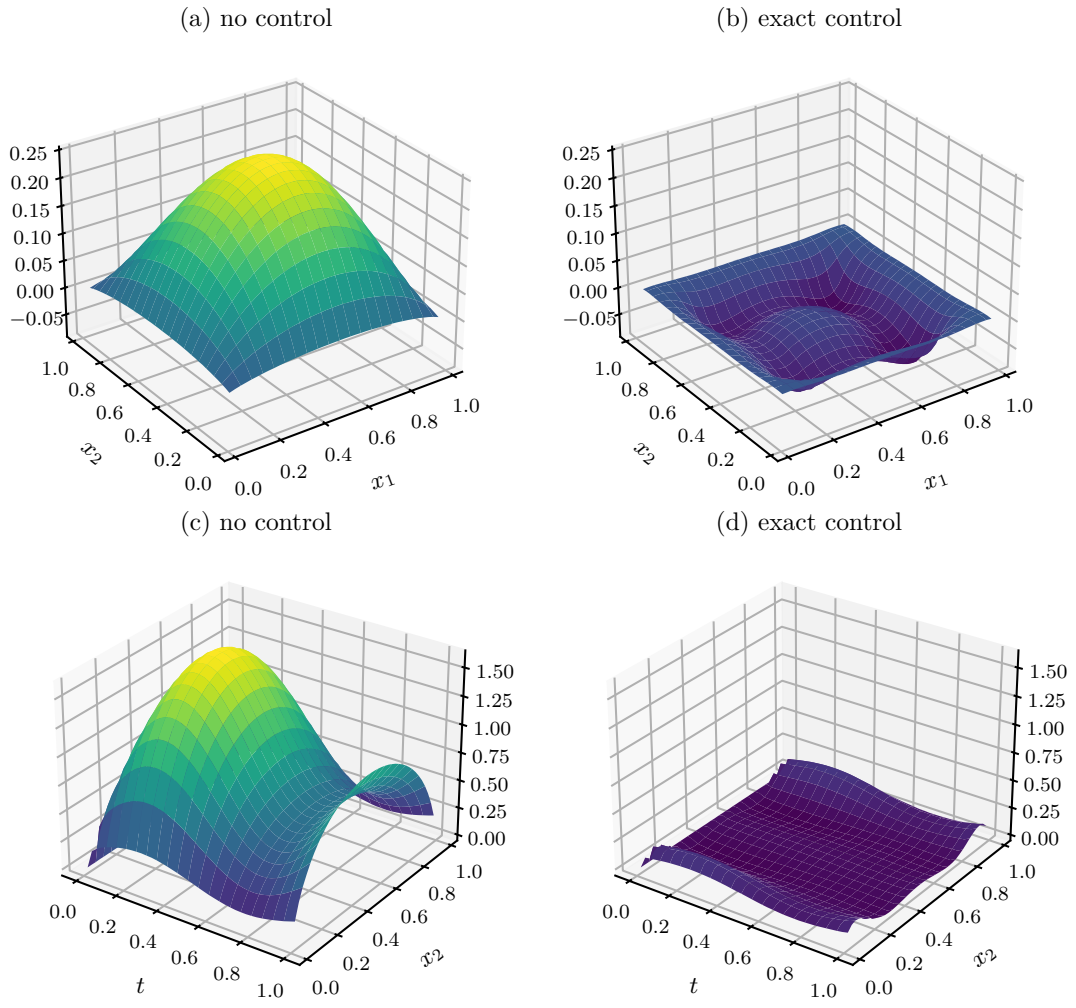


FIGURE 10. Two-dimensional habitat, $N = 40$. Top: density at $t = T/2$ integrated over the ages, $\int_0^A y(T/2, a, x)\, da$, as a function of (x_1, x_2), (a) without control and (b) with the exact control. Bottom: outgoing flux through the right side integrated over the ages, $-\int_0^A \partial_n y(t, a, 1, x_2)\, da$, as a function of (t, x_2), (c) without control and (d) with the exact control. Panels (a)–(b) have the same vertical scale, and panels (c)–(d) also.

8. Conclusions

We computed insensitizing controls with respect to domain variations for an age-structured population model with external recruitment, in one and two space dimensions. The sensitivity is a boundary integral of the product of the normal derivatives of the state and of an adjoint state. When the source is nonnegative, the sensitivity is nonnegative, and in dimension one it vanishes only if the source vanishes. It is also monotone with respect to the recruitment, the mortality and the observation region. For this reason, when the recruitment is not confined to the control region, exact insensitizing needs controls for which the source takes negative values. Such controls make sense for deviations from a reference population. We used a scheme along the characteristic lines with exact discrete gradients, a penalized least-squares method and a Newton correction in a space of small dimension. With these tools, we obtain exact insensitizing controls for a finite number of deformation modes. The cost is a few hundred solutions of the state and adjoint equations. We proved the stability of the scheme and the convergence of the discrete penalized controls. Under a criterion that can be computed, we also proved the convergence of the discrete exact controls. In the computations, the controls reduce the variation of the observation under finite deformations by two orders of magnitude. Their cost depends mainly on the size and the position of the control region, and only a little on the mortality and the diffusion. The control region should surround the monitored region.

Several questions are still open. We do not have error estimates with rates for the scheme, nor a proof of convergence of the discrete Hadamard formula. The existence and the convergence of exact controls of minimal norm are also open. In dimension two, the choice of the deformation modes and the cost of approximate insensitizing in all directions need more study. For models with a renewal law, the condition at age zero couples the characteristic lines. The algorithms can still be used, but the transposed recursions must then be computed time step by time step.

Data availability

The Python code used for the computations is available from the corresponding author on reasonable request. All the data and parameters needed to reproduce the computations are given in Sections 6 and 7.

Conflicts of interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

References

[1] Adepoju, J., Olubanwo, O., Ajani, A., & Idowu, S. (2024). Solution of Fisher's equation using homotopy perturbation Aboodh transform method. *Annals of Mathematics and Computer Science*, 23, 48-59. https://doi.org/10.56947/amcs.v23.301

[2] Ainseba, B. (2002). Exact and approximate controllability of the age and space population dynamics structured model. *Journal of Mathematical Analysis and Applications*, 275(2), 562-574. https://doi.org/10.1016/S0022-247X(02)00238-X

[3] Ainseba, B., Echarroudi, Y., & Maniar, L. (2013). Null controllability of a population dynamics with degenerate diffusion. *Differential and Integral Equations*, 26(11/12), 1397-1410. https://doi.org/10.57262/die/1378327432

[4] Al-Rawdanee, E. H., & Al-Hawasy, J. (2026). The classical continuous optimal control of coupled fourth order linear parabolic equations. *Gulf Journal of Mathematics*, 22(2). https://doi.org/10.56947/gjom.v22i2.4208

[5] Aniţa, S. (2000). *Analysis and control of age-dependent population dynamics*. Kluwer Academic Publishers, Dordrecht. https://doi.org/10.1007/978-94-015-9436-3

[6] Bahloul, R., & Rachad, H. (2026). Controllability of variable-order conformable systems. *Annals of Mathematics and Computer Science*, 36. https://doi.org/10.56947/amcs.v36.905

[7] Bodart, O., & Fabre, C. (1995). Controls insensitizing the norm of the solution of a semilinear heat equation. *Journal of Mathematical Analysis and Applications*, 195(3), 658-683. https://doi.org/10.1006/jmaa.1995.1382

[8] Boyer, F. (2013). On the penalised HUM approach and its applications to the numerical approximation of null-controls for parabolic problems. *ESAIM: Proceedings*, 41, 15-58. https://doi.org/10.1051/proc/201341002

[9] Boyer, F., Hernández-Santamaría, V., & de Teresa, L. (2019). Insensitizing controls for a semilinear parabolic equation: a numerical approach. *Mathematical Control and Related Fields*, 9(1), 117-158. https://doi.org/10.3934/mcrf.2019007

[10] Bugariu, I. F., & Micu, S. (2014). A numerical method for the controls of the heat equation. *Mathematical Modelling of Natural Phenomena*, 9(4), 65-87. https://doi.org/10.1051/mmnp/20149405

[11] Dekhkonov, F., Usubjonov, A., & Mukhammadiyev, F. (2026). Time-optimal control of the heat conduction equation. *Gulf Journal of Mathematics*, 23(2). https://doi.org/10.56947/abynw534

[12] Ervedoza, S., Lissy, P., & Privat, Y. (2022). Desensitizing control for the heat equation with respect to domain variations. *Journal de l'École polytechnique, Mathématiques*, 9, 1397-1429. https://doi.org/10.5802/jep.209

[13] Fernández-Cara, E., & Münch, A. (2012). Numerical null controllability of semi-linear 1-D heat equations: fixed point, least squares and Newton methods. *Mathematical Control and Related Fields*, 2(3), 217-246. https://doi.org/10.3934/mcrf.2012.2.217

[14] Glowinski, R., Lions, J.-L., & He, J. (2008). *Exact and approximate controllability for distributed parameter systems: a numerical approach*. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511721595

[15] Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM, Philadelphia. https://doi.org/10.1137/1.9780898719697

[16] Kamolov, S. (2026). Surrogate-accelerated Landweber iteration for stiffness identification. *Annals of Mathematics and Computer Science*, 36. https://doi.org/10.56947/amcs.v36.979

[17] Ladyženskaja, O. A., Solonnikov, V. A., & Ural'ceva, N. N. (1968). *Linear and quasilinear equations of parabolic type*. American Mathematical Society, Providence, RI. https://doi.org/10.1090/mmono/023

[18] Lemoine, J., Marín-Gayte, I., & Münch, A. (2021). Approximation of null controls for semilinear heat equations using a least-squares approach. *ESAIM: Control, Optimisation and Calculus of Variations*, 27, 63. https://doi.org/10.1051/cocv/2021062

[19] Lissy, P., Privat, Y., & Simporé, Y. (2019). Insensitizing control for linear and semilinear heat equations with partially unknown domain. *ESAIM: Control, Optimisation and Calculus of Variations*, 25, 50. https://doi.org/10.1051/cocv/2018035

[20] Maity, D., Tucsnak, M., & Zuazua, E. (2019). Controllability and positivity constraints in population dynamics with age structuring and diffusion. *Journal de Mathématiques Pures et Appliquées*, 129, 153-179. https://doi.org/10.1016/j.matpur.2018.12.006

[21] Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer, New York. https://doi.org/10.1007/978-0-387-40065-5

[22] Protter, M. H., & Weinberger, H. F. (1984). *Maximum principles in differential equations*. Springer, New York. https://doi.org/10.1007/978-1-4612-5282-5

[23] Soma, M., & Simporé, Y. (2026). Insensitizing controls with respect to domain variations for age-structured population dynamics with renewal. Manuscript submitted for publication.

[24] Squire, W., & Trapp, G. (1998). Using complex variables to estimate derivatives of real functions. *SIAM Review*, 40(1), 110-112. https://doi.org/10.1137/S003614459631241X

[1] Département de Mathématiques, Université Thomas SANKARA, Centre Universitaire de Tenkodogo, 12 BP 417 Ouagadougou 12, Burkina Faso.
Email address: mifiambasoma@yahoo.fr

[2] École Normale Supérieure, Ouagadougou, Burkina Faso.
Email address: aliouedraogo95@yahoo.fr, ouedraogom40@yahoo.fr

[3] Classes Préparatoires aux Études d'Ingénieur, École Polytechnique de Ouagadougou, 08 BP 143 Ouagadougou 08, Burkina Faso.
Email address: zongoyac10@gmail.com