Picard Iteration for Parameter Estimation in Nonlinear Ordinary Differential Equations

Aleksandr Talitckii [a], Matthew M. Peet [a],

[a] *School for the Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, 85298 USA*

Abstract

We consider the problem of using experimental time-series data for parameter estimation in nonlinear ordinary differential equations, focusing on the case where the data is noisy, sparse, irregularly sampled, includes multiple experiments, and does not directly measure the system state or its time-derivative. To account for such low-quality data, we propose a new framework for gradient-based parameter estimation which uses the Picard operator to reformulate the problem as constrained optimization with infinite-dimensional variables and constraints. We then use the contractive properties of the Picard operator to propose a class of gradient-contractive algorithms and provide conditions under which such algorithms are guaranteed to converge to a local optima. The algorithms are then tested on a battery of models and variety of datasets in order to demonstrate robustness and improvement over alternative approaches.

Key words: parameter estimation, system identification, nonlinear systems, gradient-based methods, optimization algorithms

1 Introduction

Mechanical, electrical, chemical and biological processes are often dynamic with highly nonlinear response to disturbance, actuation, and initial state. Regulation of such processes requires accurate models of this nonlinear behaviour. Ideally, the dynamics of the process can be modelled through the combination of well-established physical principles. In some cases, however (e.g. biological systems), the physical principles governing the behaviour are speculative, over-simplified, or only partially understood. Furthermore, even when the physics of the problem are well-understood, there may still be substantial uncertainty in parameters of the model due to natural variation, errors in measurement, oversimplification of the physics or difficulty in direct measurement of these parameters. In such cases, there may be only a few uncertain parameters in the model, and they may be constrained to lie in some set (e.g. positivity is a common form of parameter constraint).

When it is not possible to measure the parameters of the system directly, these parameters must be inferred indirectly by examination of the response of the system to variations in input and initial state [1, 2]. In the extreme case, where there is no understanding of the physics, we have black-box system identification [3] — e.g. Koopman operators [4–7]; extended dynamic mode decomposition (EDMD) [8]; sparse identification of nonlinear dynamics (SINDy) [9–12]; autoregressive models [13]; and neural networks [14].

When the physics are only partially understood or the system parameters are constrained, however, we have grey-box system identification [15] – a more challenging problem than black-box modelling. Specifically, suppose the system to be identified is known to have the form

$$\begin{aligned} \dot{x}(t) &= f(t, x(t), \theta) \qquad x(0) = x_0 \\ y(t) &= g(t, x(t), \theta) \end{aligned} \tag{1}$$

where the functions f and g are given, $x(t) \in \mathbb{R}^n$ is the internal state, x_0 is the initial state, and $y(t) \in \mathbb{R}^{n_y}$ is some measurable output. The vector, $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$, represents the unknown parameters in the model. The dependence of f and g on time, t, is typically used to represent the effect of known inputs, $u(t)$ — in which case we might equivalently write the functions as $f(x(t), u(t), \theta)$ and $g(x(t), u(t), \theta)$. In the parameter estimation problem, we are commonly given several input-output pairs [16], $u_i(t), y_i(t)$ so in this case, each input would require a different f and g — i.e. $f(x(t), u_i(t), \theta)$ and $g(x(t), u_i(t), \theta)$. Implicit in the use of measured outputs, $y(t)$, is that the initial condition, x_0 is not directly measurable and is likewise different for every experiment. This initial condition must then be estimated along with the unknown

⋆ his work was supported in part by the National Science Foundation under Grants NSF CCF-2323532 and NSF/NIH 2054354 awarded to M. M. Peet. (Corresponding author: Aleksandr Talitckii.)

Email addresses: atalitck@asu.edu (Aleksandr Talitckii), mpeet@asu.edu (Matthew M. Peet).

parameters (similar to the inverse problem [17]).

For a given parameterized vector field, $f(t, x, \theta)$, we suppose that Eqn. (1) is well-posed and define the corresponding solution map, $\phi_f(t, x, \theta)$, as the unique function which satisfies

$$\partial_t \phi_f(t, x, \theta) = f(t, \phi_f(t, x, \theta), \theta), \quad \phi_f(0, x, \theta) = x$$

for all $t \geq 0$, $x \in \mathbb{R}^{n_x}$ and $\theta \in \Theta$. While it is almost never possible to obtain an analytic expression for the solution map of a nonlinear system, the ϕ_f notation allows us to efficiently represent the problem of data-based system identification. Specifically, suppose that experimental data is available from solutions of Eqn. (1) consisting of measurements y_i at a sequence of discrete times $\{t_i \in \mathbb{R}\}_{i=1}^{N_s}$ so that $y_i = (1+n_i)g(t_i, \phi_f(t_i, x, \theta), \theta) + m_i$ where θ, x are unknown and $m_i \in \mathbb{R}^{n_y}$ and $n_i \in \mathbb{R}$ represent measurement errors. We denote the set of all such measurements as $Y = \{(y_i, t_i) \in \mathbb{R}^{n_y} \times \mathbb{R} \ : \ i = 1, ..., N_s\}$.

We may now formulate the parameter estimation/system identification problem as

$$\min_{\theta \in \Theta, x \in \mathbb{R}^n} L_{ls}(x, \theta, \phi_f(\cdot, x, \theta)), \tag{2}$$

where for simplicity we use the least-squares *Loss Function* [18] defined as

$$L_{ls}(x, \theta, \phi_f(\cdot, x, \theta)) = \frac{1}{2N_s} \sum_{i=1}^{N_s} \|y_i - g(t_i, \phi_f(t_i, x, \theta), \theta)\|_2^2.$$

Problem (2) is an unconstrained, typically non-convex optimization problem [19] in parameters x, θ. One approach to solving such a problem is to use numerical simulation in place of the solution map [20] and then apply gradient-free, black box optimization tools such as numerical gradient-approximation [21,22], parameter cascade [23, 24], nonparametric estimators [25–27], genetic algorithms [28, 29], and simulated annealing [30]. Unfortunately, however, such methods are computationally inefficient and may fail to converge to even locally optimal solutions.

By contrast, gradient-based methods are guaranteed to converge to local optima but require some method for computing the gradient of the solution map, $\nabla_{x,\theta}\phi_f$. Because analytic expressions for the solution map are rarely available, however, such a gradient descent approach is not practical.

Faced with the limitations of black-box optimization and lack of analytic solution maps, and motivated by the use of gradient descent to obtain local optima, we consider now an alternative formulation of the parameter estimation/system identification problem posed in Eqn. (2), which does not require knowledge of the solution map. Specifically, in Sec. 2, we reformulate the unconstrained optimization problem as the equivalent constrained optimization problem defined as

$$\begin{aligned} \min_{x \in X, \theta \in \Theta, \mathbf{u} \in C(\mathbb{R})} & L_{ls}(x, \theta, \mathbf{u}) \\ \text{s.t. } & \mathbf{u}(t) = x + \int_0^t f(s, \mathbf{u}(s), \theta) ds \qquad \forall t \geq 0. \end{aligned} \tag{3}$$

In Optimization Problem (3), we have replaced the solution map ϕ_f by the variable $\mathbf{u}$ and added the constraint that $\mathbf{u}$ be a solution of the given system. Clearly, Problems (2) and (3) are equivalent. Moreover, the constraint, $\mathbf{u}(t) = x + \int_0^t f(s, \mathbf{u}(s), \theta) ds$, may be conveniently represented using the Picard operator as $\mathbf{u} = \mathcal{P}_{x,\theta}\mathbf{u}$, where

$$(\mathcal{P}_{x,\theta}\mathbf{u})(t) := x + \int_0^t f(s, \mathbf{u}(s), \theta) ds.$$

Although Optimization Problem (3) does not require the solution map, it introduces an infinite-dimensional variable $\mathbf{u}$ and associated equality constraint which must hold for all times, t, on which the data is defined. Solving such problems without an explicit representation of the feasible set is known to be hard [31].

To address these challenges, in Sec. 3 we propose a new class of gradient-contractive algorithms (Alg. 3) for solving the formulation of the parameter estimation problem in Eqn. (3). In this algorithm, each iteration has two steps – gradient computation and contraction to the feasible set. Specifically, the first step computes the gradient of the loss function and updates the initial condition and parameter variables as x_{k+1}, θ_{k+1} where

$$\begin{aligned} x_{k+1} &= x_k - \alpha \nabla_x L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_k) \\ \theta_{k+1} &= \theta_k - \alpha \nabla_\theta L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_k), \end{aligned}$$

and where the n^{th} order Picard iteration, $\mathcal{P}^n_{x,\theta}$, is introduced into the objective so that the gradient accounts for multipliers corresponding to the KKT conditions for the constraint $\mathbf{u} = \mathcal{P}_{x,\theta}\mathbf{u}$ and where the step size $\alpha > 0$ is computed through a line search. The second step contracts the variable $\mathbf{u}_k(t)$ to the new feasible set (determined by x_{k+1} and θ_{k+1}) using the contractive property of the Picard iteration as

$$\mathbf{u}_{k+1}(t) = (1 - \sigma)\mathbf{u}_k(t) + \sigma(\mathcal{P}_{x_{k+1},\theta_{k+1}}\mathbf{u}_k)(t)$$

for some $\sigma \in (0, 1]$. This second step translates the solution closer to the infinite-dimensional manifold described by the feasible set of Problem (3) – imitating the projection step of gradient-projection algorithms and resolving the problems introduced by the infinite-dimensional nature of the equality constraints. This approach also eliminates the need to explicitly parameterize the infinite-dimensional variable, $\mathbf{u}$. A limitation of this approach, of course, is the need for the Picard operator to be contractive on the interval of time on which the data was sampled. However, as described in Sec. 3, an extended form of Picard operator may be used which is contractive over arbitrary intervals (assuming well-posedness of the solution map).

Note that Picard iteration has previously been used for analysis of black-box system identification [32], and parameter estimation in [33,34]. Specifically, in [33] gradient descent of the unconstrained Problem (2) was considered for signal transduction networks, wherein a Picard iterate was used in lieu of the solution map (See Alg. 1). Unfortunately, convergence of the Picard iteration is typically limited to short time intervals, thus limiting the generality of this approach. An attempt to extend this approach to longer time intervals using multiple shooting methods was proposed in [34]. However, convergence of neither method has been proven or extended to sparse data, irregular sampling times or measured outputs.

In contrast to the work in [33, 34], the use of a constrained formulation of the parameter estimation problem, the use of gradient contractive algorithms, and the proposed extended form of Picard iteration (Alg. 4) used in this paper allow for establishment of convergence proofs, arbitrary sampling intervals, and measured outputs. To establish these results, in Sec. 4, we first consider the general class of two-step gradient contractive algorithms and establish conditions for convergence based on properties of the contractive operator, $\mathcal{P}$. Next, in Sec. 5, we examine the Picard operator $\mathcal{P}_{x,\theta}$ and establish conditions for convergence to the solution map and its gradient. In Sec. 6, we show these results apply to the extended Picard operator (Alg. 4) and provide conditions for convergence of Alg. 4 to a local solution of the parameter estimation problem. The results are then applied to a rigorous battery of numerical tests, including: Van der Pol oscillator to evaluate the effect of the regularization parameter; the FitzHugh-Nagumo Neuron to evaluate the effect of irregular sampling; the Rosenzweig-MacArthur predator-prey model to evaluate the effect of sparse data; a tumor growth model to investigate the question of identifiability; and the Lorentz system for comparison with SINDy and black-box system optimization methods.

Notation

We use $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_+$, to denote the natural numbers, the real numbers, and the non-negative real numbers, respectively. For a given $J \in \mathbb{N}$ we denote $\overline{1,J} = \{j \in \mathbb{N} \mid 1 \leq j \leq J\}$. For $a, b \in \mathbb{R}_+$ we denote $\lfloor a/b \rfloor := \underset{m\in\mathbb{N}, a\geq mb}{\texttt{argmax}}\ m$ and $a \bmod b := a - \lfloor a/b \rfloor b$. For a given sets X, Y we denote $\mathcal{F}(X, Y)$ to be the set of all functions $f : X \to Y$. For compact $S \subset \mathbb{R}^n$, we denote $C^k(S)$ to be the space of k times continuously differentiable functions, $f \in \mathcal{F}(S, \mathbb{R}^m)$ with norm $\|f\| = \sup_{s\in S} \|f(s)\|_2$. We use $C_a(S) \subset C(S) := C^0(S)$ to denote the ball of radius a as $C_a(S) = \{\mathbf{u} \in C(S) \mid \|\mathbf{u}\| \leq a\}$ and use C_a when the domain is clear from context. For $u \in C^1(S)$, we use $\nabla_x u$ to denote the gradient as $\nabla_x u = \left[\frac{\partial}{\partial x_1} u, \ldots, \frac{\partial}{\partial x_m} u\right]^T$. For compact $X \subseteq \mathbb{R}^n$ we denote $\Pi_X : \mathbb{R}^n \to X$ to be the projection operator so that $\Pi_X(y) = \underset{x\in X}{\texttt{argmin}} \|x - y\|_2$. We say $f \in \mathcal{F}(X, \mathbb{R})$ is strongly convex on $X \subset \mathbb{R}^n$ with modulus $\mu \geq 0$ if $f(x) - \frac{\mu}{2}\|x\|^2$ is convex on X. We say $f \in \mathcal{F}(X \times Y, \mathbb{R})$ is strongly convex on $X \subset \mathbb{R}^n$ uniformly in y if there exists $\mu \geq 0$ such that $f(x, y) - \frac{\mu}{2}\|x\|^2$ is convex on X for all fixed $y \in Y \subset \mathbb{R}^m$. For $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$, we say that $f \in \mathcal{F}(X, Y)$ is Lipschitz continuous with constant K if for all $x_1, x_2 \in X$ we have $\|f(x_1) - f(x_2)\| \leq K\|x_1 - x_2\|$ and for $Z \subset \mathbb{R}^p$, $f \in \mathcal{F}(X \times Z, Y)$ is Lipschitz continuous with respect to x with constant K if for all $x_1, x_2 \in X$ and $z \in Z$ we have $\|f(x_1, z) - f(x_2, z)\| \leq K\|x_1 - x_2\|$.

2 Formulating the Parameter Estimation Problem for Nonlinear Differential Equations

Consider a parameterized nonlinear differential equation model in state-space representation.

$$\begin{aligned} \dot{\mathbf{u}}(t) &= f(t, \mathbf{u}(t), \theta) \qquad \mathbf{u}(0) = x \\ y(t) &= g(t, \mathbf{u}(t), \theta) \end{aligned} \tag{4}$$

where $t \in \Gamma$ is time (on interval Γ), $\mathbf{u}(t) \in \mathbb{R}^{n_x}$ is system state, $\theta \in \Theta$ are unknown parameters (restricted to compact convex $\Theta \subset \mathbb{R}^{n_\theta}$), x are unknown initial conditions (restricted to compact convex $X \subset \mathbb{R}^{n_x}$), and $y(t) \in \mathbb{R}^{n_y}$ are measured outputs. The function $f \in C^1(\Gamma \times \mathbb{R}^{n_x} \times \Theta)$ is the vector field and $g \in C^1(\Gamma \times \mathbb{R}^{n_x} \times \Theta)$ is the state-to-output map. We assume existence and uniqueness, and continuous dependence of solutions on initial conditions and parameters, implying existence of a solution map, $\phi : \Gamma \times X \times \Theta \to \mathbb{R}^{n_x}$, which satisfies

$$\partial_t \phi(t, x, \theta) = f(t, \phi(t, x, \theta), \theta), \quad \phi(0, x, \theta) = x. \tag{5}$$

Now suppose we are given a set of measurements of the dynamical system for some initial condition x and true parameter value, θ as

$$Y = \{(y_i, t_i) | y_i = (1 + n_i) g(t_i, \phi(t_i, x, \theta), \theta) + m_i, i \in \overline{1, N_s}\},$$

where $m_i \in \mathbb{R}^{n_y}$ and $n_i \in \mathbb{R}$ represent measurement errors.

The parameter estimation problem is to find initial condition, $x \in X$, and parameter values, $\theta \in \Theta$, that minimize the least squares error in the predicted output relative to the set of measurements. This can be formulated as either an unconstrained or a constrained optimization problem depending on whether we know the solution map. If the solution map, ϕ, is known, the problem can be formulated as an unconstrained optimization problem of the form

$$\min_{x\in X, \theta\in\Theta} L_{ls}(x, \theta, \phi(\cdot, x, \theta)), \tag{6}$$

where

$$L_{ls}(x, \theta, \phi(\cdot, x, \theta)) = \frac{1}{2N_s} \sum_{i=1}^{N_s} \|y_i - g(t_i, \phi(t_i, x, \theta), \theta)\|^2.$$

Since the solution map is unknown, however, we may reformulate the parameter estimation problem by instead introducing a variable, $\mathbf{u}$, representing the solution of the ODE and add the constraint that this variable be a solution of the ODE.

$$\begin{aligned} &\min_{x\in X, \theta\in\Theta, \mathbf{u}\in C(\Gamma)} L_{ls}(x, \theta, \mathbf{u}) \\ &\qquad \text{s.t. } \mathbf{u}(t) = x + \int_0^t f(s, \mathbf{u}(s), \theta) ds, \end{aligned} \tag{7}$$

where the integral constraint is a necessary and sufficient condition for $\mathbf{u}(t)$ to satisfy the ODE (Eqn. (4)) with initial condition $\mathbf{u}(0) = x$ and parameter values θ. The advantage of this formulation is that it does not require knowledge of the solution map. However, the disadvantage is that we have introduced an infinite-dimensional optimization variable in the form of $\mathbf{u}$.

2.1 Reformulation using the Picard Operator

The unconstrained optimization problem in Eqn. (6) is expressed in terms of the solution map (which is unknown) and the constrained optimization problem in Eqn. (7) is expressed in terms of the variable $\mathbf{u}$ (which is infinite-dimensional variable and where the constraint must hold at an infinite number of times). The problems with both these formulations may be ameliorated through use of what is known as the Picard operator, $\mathcal{P}_{x,\theta}$.

Definition 1 *For a given $t_0 \in \Gamma$, $T > 0$ and continuous vector field, $f : \Gamma \times \mathbb{R}^{n_x} \times \Theta \to \mathbb{R}^{n_x}$, we define the Picard operator $\mathcal{P}_{t_0,x,\theta} : C[0,T] \to C[0,T]$*

$$(\mathcal{P}_{t_0,x,\theta}\mathbf{u})(t) := x + \int_0^t f(s + t_0, \mathbf{u}(s), \theta)ds. \qquad (8)$$

and when t_0=0, we simplify the notation as $\mathcal{P}_{x,\theta}$:=$\mathcal{P}_{0,x,\theta}$.

The critical property of the Picard operator is that if T is sufficiently small, iterations of the Picard operator converges to the solution of $\dot{x} = f(x)$. That is, for any $\mathbf{u}(t)$, $\mathcal{P}_{x,\theta} \circ \cdots \circ \mathcal{P}_{x,\theta}\mathbf{u} \to \mathbf{u}^*$ where $\mathcal{P}_{x,\theta}\mathbf{u}^* = \mathbf{u}^*$.

The Picard operator allows us to *approximate* Optimization Problem (6) for fixed $n \in \mathbb{N}$ as

$$\min_{x\in X,\theta\in\Theta} L_{ls}(x, \theta, \mathcal{P}^n_{x,\theta}\mathbf{u}), \qquad (9)$$

and to *restate* Optimization Problem (7) as

$$\min_{\substack{x\in X,\theta\in\Theta,\\ \mathbf{u}\in C(\Gamma)}} L_{ls}(x, \theta, \mathbf{u}) \quad \text{s.t.} \quad \mathcal{P}_{x,\theta}\mathbf{u} = \mathbf{u}. \qquad (10)$$

Now, the approximation of unconstrained Optimization Problem (6) in Problem (9) has eliminated the need for knowledge of the solution map, but introduced a very complicated dependence on initial state, x and parameter, θ. Meanwhile, the restatement of constrained Optimization Problem (7) in (10) does not seem to have any immediate benefits. As will be discussed in the following section, however, the contractive properties of the solution map will allow us to eliminate the need for infinite-dimensional variables and constraints in (10).

3 Gradient Descent for Parameter Estimation

In this section we consider algorithms to solve the optimization problems in Sec. 2. For the unconstrained problem (Problems (6) and (9)), we focus on calculating the gradient of high-order Picard iterations. For the constrained optimization problem (Problems (7) and (10)), we focus on an iteration which eliminates the infinite-dimensional nature of the constraint and variables.

3.1 A Gradient Descent Algorithm for the Unconstrained Formulation

As indicated in Subsec. 2.1, Optimization Problem (6) can be approximated using the Picard iteration as in (9), where the Picard operator in this case is understood to act on the set of functions $\mathbf{u}(t, x, \theta)$ and convergence of Picard iterates is then to the solution map. Unconstrained optimization problems of this form can be solved using gradient descent (combined with projection onto the convex set of allowable parameters, $\theta \in \Theta$ and initial conditions, $x \in X$). Specifically, let $\alpha > 0$ be a step size and recall that the projection operator $\Pi_X : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ is defined as

$$\Pi_X(z) := \texttt{argmin}_{x\in X} \|z - x\|_2^2 \qquad (11)$$

We may now propose Alg. 1, where the gradient is computed using the n-th order Picard iterate approximation of the solution map. As stated in the following theorem, this gradient also converges to the gradient of the solution map – implying that Alg. 1 is actually an approximation of the gradient descent algorithm as applied the original formulation of the problem in Eqn. (6).

Algorithm 1 Gradient Descent Algorithm

`INPUT:` N – the number of iterations
n – the order of Picard iterations
$x_0 \in X$, $\theta_0 \in \Theta$, $\mathbf{u}_0 = x_0$, $\alpha > 0, \sigma \in (0, 1]$
`For` k from 0 to $N - 1$
$\theta_{k+1} := \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_0)]$
$x_{k+1} := \Pi_X[x_k - \alpha\nabla_x L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_0)]$
`EndFor`
`OUTPUT:` x_N, θ_N

Algorithm 2 Gradient-Contractive Algorithm

`INPUT:` N – the number of iterations
$x_0 \in X$, $\theta_0 \in \Theta$, $\mathbf{u}_0 = x_0$, $\alpha > 0, \sigma \in (0, 1]$
`For` k from 0 to $N - 1$
$\theta_{k+1} := \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{ls}(x_k, \theta_k, \mathbf{u}_k)]$
$x_{k+1} := \Pi_X[x_k - \alpha\nabla_x L_{ls}(x_k, \theta_k, \mathbf{u}_k)]$
$\mathbf{u}_{k+1}(t) := (1 - \sigma)\mathbf{u}_k(t) + \sigma(\mathcal{P}_{x_{k+1},\theta_{k+1}}\mathbf{u}_k)(t)$
`EndFor`
`OUTPUT:` $x_N, \theta_N, \mathbf{u}_N$

Algorithm 3 Gradient-Contract-Multiplier Algorithm

`INPUT:` N – the number of iterations
n – the order of Picard iterations
$x_0 \in X$, $\theta_0 \in \Theta$, $\mathbf{u}_0 = x_0$, $\alpha > 0, \sigma \in (0, 1]$
`For` k from 0 to $N - 1$
$\theta_{k+1} := \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_k))]$
$x_{k+1} := \Pi_X[x_k - \alpha\nabla_x L_{ls}(x_k, \theta_k, \mathcal{P}^n_{x_k,\theta_k}\mathbf{u}_k)]$
$\mathbf{u}_{k+1}(t) := (1 - \sigma)\mathbf{u}_k(t) + \sigma(\mathcal{P}_{x_{k+1},\theta_{k+1}}\mathbf{u}_k)(t)$
`EndFor`
`OUTPUT:` $x_N, \theta_N, \mathbf{u}_N$

Algorithm 4 Extended Grad-Contract-Multiplier Alg.

`INPUT:` N – the number of iterations
n – the order of Picard iterations
J – the number of time intervals
$T > 0$, $\theta_0 \in \Theta$, $x_0 \in X^{\otimes J}$, $\mathbf{u}_0 \in C^{\otimes J}[0, T)$,
$\lambda \geq 0$, $\alpha > 0, \sigma \in (0, 1]$
`INIT:` $\mathbf{u}_j(t) = x_{0,j}$ for all $j \in \overline{1, J}$
`DEFINE:` $L_{\lambda,n}(x, \theta, \mathbf{u})$ as in Eqn. (16)
`For` k from 0 to $N - 1$
$\theta_{k+1} := \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{\lambda,n}(x_k, \theta_k, \mathbf{u}_k)]$
`For` j from 1 to J
$x_{k+1,j} := \Pi_X[x_{k,j} - \alpha\nabla_{x_j} L_{\lambda,n}(x_k, \theta_k, \mathbf{u}_k)]$
$\mathbf{u}_{k+1,j} := (1 - \sigma)\mathbf{u}_{k,j} + \sigma(\mathcal{P}_{(j-1)T,x_{k+1,j},\theta_{k+1}}\mathbf{u}_{k,j})$
`EndFor`
`EndFor`
`OUTPUT:` $x_N, \theta_N, \mathbf{u}_N$

Theorem 2 *Let Γ, X, Θ be compact and T be sufficient small. Let $a = 2\sup_{x\in X}\|x\|_2$ and $f \in C^1(\Gamma \times B_a \times \Theta)$ and $\nabla_x f, \nabla_\theta f$ are Lipschitz continuous functions. Then for all $t \in [0,T]$, $x \in X$, $\theta \in \Theta$ and for any $\mathbf{u} \in C([0,T])$ such that $\|\mathbf{u}\|_\infty \le a$ we have $\lim_{n\to\infty} \nabla_{x,\theta}(\mathcal{P}^n_{x,\theta}\mathbf{u})(t) = \nabla_{x,\theta}\phi(t,x,\theta)$, where $\phi(t,x,\theta)$ is the solution map (Eqn. (5)).*

PROOF. See Prop. 14 in Sec. 5. ■

Convergence of the gradient descent in Alg. 1 to the solution of Optimization Problem (6) as $n \to \infty$ is then guaranteed if sets X, Θ are convex and the loss function is strongly convex – See, e.g., Thm. 4.32 in [35]. Of course, in reality, for most parameter estimation problems the dependence of the loss function in Eqn. (9) on the system parameters and initial state is not convex, and hence Alg. 1 is only guaranteed to converge to a set of local minima.

More significantly, convergence is only guaranteed in the limit $n \to \infty$. However, when n becomes large, computation of the gradient $\nabla_{x,\theta}\mathcal{P}^n_{x,\theta}$ becomes difficult. For this reason, we now consider an algorithm for solving the constrained form of Optimization Problem (10).

3.2 Gradient Contraction for Constrained Optimization

Now let us consider the constrained form of the optimization problem given in (10). To account for the constraint, we use a two-step gradient contractive algorithm which includes both a gradient step on variables x, θ and a contraction step on variable $\mathbf{u}$. In a standard gradient *projection* algorithm, we would evaluate the gradient with respect to $x, \theta, \mathbf{u}$ and then project $\mathbf{u}$ onto the feasible set. However, because $\mathbf{u}$ is infinite-dimensional and the feasible set is a manifold on this infinite-dimensional space, such an approach is difficult. As an alternative, however, one might propose Alg. 2, which does not explicitly parameterize the variable $\mathbf{u}$, but instead takes the gradient with respect to x and θ for some fixed $\mathbf{u}$ as

$$\theta_{k+1} = \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{ls}(x_k, \theta_k, \mathbf{u}_k)]$$
$$x_{k+1} = \Pi_X[x_k - \alpha\nabla_x L_{ls}(x_k, \theta_k, \mathbf{u}_k)],$$

where note that, unlike in Alg. 1, there is no need to compute the gradient of a Picard iteration. Now, instead of parameterizing $\mathbf{u}$ explicitly, for given x, θ, we use the Picard iteration to update our variable $\mathbf{u}$ as

$$\mathbf{u}_{k+1} = (1-\sigma)\mathbf{u}_k + \sigma\mathcal{P}_{x_{k+1},\theta_{k+1}}\mathbf{u}_k$$

for some step size, $\sigma \in (0,1]$. Then, because the Picard iteration is a contraction on sufficiently short time intervals, this step not only updates $\mathbf{u}$, but moves it closer to the feasible set $\{\mathbf{u} : \mathbf{u} = \mathcal{P}_{x_{k+1},\theta_{k+1}}\mathbf{u}\}$.

Now, Lipschitz continuity and strong convexity of the loss function with respect to x, θ implies convergence of Alg. 2 to a fixed point x_0^*, θ_0^* and $\mathbf{u}_0^*$. However, our failure to explicitly parameterize $\mathbf{u}$ does have a cost. Specifically, while the fixed point will be feasible, this fixed point will not necessarily be optimal. To see this, consider the KKT conditions necessary for optimality of Problem (10).

Proposition 3 *Suppose $f, g \in C^1(\Gamma \times \mathbb{R}^{n_x} \times \Theta)$ and $\{x, \theta, \mathbf{u}\} \in int(X) \times int(\Theta) \times C[0,T]$ is a stationary point of Optimization Problem (7). Then*

$$\nabla_{x,\theta}L_{ls}(x,\theta,\mathbf{u}) + \sum_{i=1}^{N_s}\nabla_{u_i}L_{ls}(x,\theta,\mathbf{u})\nabla_{x,\theta}\phi(t_i,x,\theta) = 0, \quad (12)$$

where $\phi(t,x,\theta)$ is the solution map of Differential Eqn. (4) and

$$\nabla_{u_i}L_{ls}(x,\theta,\mathbf{u}) := \frac{1}{N_s}(y_i - g(t_i, \mathbf{u}(t_i), \theta))^T\nabla_u g(t_i, \mathbf{u}(t_i), \theta).$$

PROOF. The proof is based on equivalence of Problems (9) and (7) and application of the KKT conditions as formulated in [36].■

Prop. 3 implies the solution, $x^*, \theta^*, \mathbf{u}^*$ of Problem (7) need satisfy Eqn. (12). However, (interior) fixed points of Alg. 2 only satisfy $\nabla_{x,\theta}L_{ls}(x,\theta,\mathbf{u}) = 0$. To resolve this issue, in the following subsection, we approximate the multiplier $\nabla_{x,\theta}\phi(t_i,x,\theta)$ in Eqn. (12) using Picard iteration and include this term in the updates of x and θ.

3.3 A Gradient Contraction Multiplier Algorithm for the Constrained Formulation

We now present Alg. 3, which does not require the use of infinite-dimensional variables and constraints, but for which the fixed point satisfies an approximation of the KKT conditions for optimality defined in Eqn. (12) of Prop. 3. Specifically, we require a fixed point $\{x, \theta, \mathbf{u}\}$ of Alg. 3 to satisfy

$$\nabla_{x,\theta}L_{ls}(x,\theta,\mathcal{P}^n_{x,\theta}\mathbf{u}) = \nabla_{x,\theta}L_{ls}(x,\theta,\mathbf{u}) \quad (13)$$
$$+ \nabla_{u(t_i)}L_{ls}(x,\theta,\mathbf{u})\nabla_{x,\theta}(\mathcal{P}^n_{x,\theta}\mathbf{u})(t_i,x,\theta) = 0$$

where recall $\lim_{n\to\infty}\nabla_{x,\theta}(\mathcal{P}^n_{x,\theta}\mathbf{u})(t) = \nabla_{x,\theta}\phi(t,x,\theta)$ as in Thm. 2. Thus for any n, the optimality conditions of Prop. 3 are satisfied in some approximate sense, where the accuracy of this approximation increases with n. As will be shown in Sec. 4 this allows us to prove convergence of the algorithm to optimality under suitable regularity and convexity conditions. As a practical matter, as will be shown in the numerical examples (Sec. 7), Alg. 3 requires only $n = 1$ or $n = 2$ for highly accurate solutions.

3.4 An Extended Gradient Contraction Multiplier Algorithm for the Constrained Formulation

Alg. 2 and Alg. 3 eliminated the need for explicit parametrization of the variable $\mathbf{u}$ by initializing $\mathbf{u}$ with some arbitrarily chosen initial guess and then using the Picard operator in lieu of a projection to enforce the equality constraint $\mathcal{P}_{x,\theta}\mathbf{u} - \mathbf{u} = 0$. However, this approach is premised on the assumption that the Picard operator is sufficiently contractive so that, after a sufficient number of iterations, this equality constraint will be satisfied. However, the time interval for which the Picard iteration is contractive is substantially smaller than the interval on which the data is typically sampled or for which the solution map is defined. As a final step, therefore, we partition the time-domain into disjoint intervals and apply the Picard iteration to each interval, using a regularization term to ensure that discontinuities between intervals are minimized.

Specifically, for a given fixed interval of convergence, $t \in [0, T]$, we lift the space of initial conditions as $x \in X^{\otimes J} := X \times \cdots \times X$ and the space of solutions as $\mathbf{u} \in C[0,T]^{\otimes J}$ so that $\mathbf{u}_j(0) = x_j$ and $\mathbf{u}_j(T) = x_{j+1}$ for all $j \in \overline{1, J-1}$ and $\mathbf{u}_j(t)$ represents the estimated solution on interval $t \in [(j-1)T, jT)$. Then the constrained optimization problem (Eqn. (7)) may now be formulated as

$$\min_{\substack{x \in X^{\otimes J}, \theta \in \Theta \\ \mathbf{u} \in C[0,T]^{\otimes J}}} \frac{1}{2N_s} \sum_{i=1}^{N_s} \|y_i - g(t_i, \mathbf{u}_{\lfloor t_i/T \rfloor}(t_i \bmod T), \theta)\|_2^2$$
$$\text{s.t. } \mathcal{P}_{(j-1)T, x_j, \theta} \mathbf{u}_j = \mathbf{u}_j \qquad \forall j,\ t \in [0,T] \qquad (14)$$
$$\mathbf{u}_j(T) = x_{j+1} \qquad \forall j \in \overline{1, J-1}.$$

To see that Problems (10) and (14) are equivalent, we have the following.

Lemma 4 *Let $\Gamma = [0, JT]$, then Optimization Problems (10) and (14) are equivalent.*

PROOF. Given a solution $\mathbf{u} \in C[0, JT], x \in X, \theta \in \Theta$ to Problem (10) with objective value γ, let $x_j = \mathbf{u}((j-1)T)$ and $\mathbf{u}_j(t) = \mathbf{u}(t + (j-1)T)$ for all $j \in \overline{1, J}$ and $t \in [0, T]$. Since $\mathbf{u}$ is a solution of Problem (10), we have $x_j = \mathbf{u}((j-1)T) = (\mathcal{P}_{0,x,\theta}\mathbf{u})((j-1)T)$ – implying $\mathbf{u}_j(T) = x_{j+1}$. Next, since $\mathcal{P}_{0,x,\theta}\mathbf{u} = \mathbf{u}$, we have

$$\begin{aligned}(\mathcal{P}_{(j-1)T,x_j,\theta}\mathbf{u}_j)(t) &= (\mathcal{P}_{(j-1)T,0,\theta}\mathbf{u}_j)(t) + x_j \\ &= (\mathcal{P}_{(j-1)T,0,\theta}\mathbf{u}_j)(t) + (\mathcal{P}_{0,x,\theta}\mathbf{u})((j-1)T) \\ &= (\mathcal{P}_{0,x,\theta}\mathbf{u})(t + (j-1)T) = \mathbf{u}_j(t)\end{aligned}$$

for all $j \in \overline{1, J}$ and $t \in [0, T]$. Thus, $\mathbf{u}_j, x_j, \theta$ is feasible for Problem (14) with objective value γ.

Conversely, given a solution $\mathbf{u}_j, x$ to Problem (14) with objective value $\hat{\gamma}$, let $\hat{\mathbf{u}}(t) = \mathbf{u}_{\lfloor t/T \rfloor}(t \bmod T)$ and $\hat{x} = x_1$. Then, since $\mathbf{u}_j(T) = x_{j+1} = \mathbf{u}_{j+1}(0)$, $\hat{\mathbf{u}}$ is continuous. Now, suppose for some $j_0 \in \overline{1, J}$ and any $t \le j_0 T$ we have $(\mathcal{P}_{0,x_1,\theta}\hat{\mathbf{u}})(t) = \hat{\mathbf{u}}(t)$, which clearly holds for $j_0 = 1$. Then, for all $t \in [j_0 T, (j_0+1)T]$ we have

$$\begin{aligned}(\mathcal{P}_{0,x_1,\theta}\hat{\mathbf{u}})(t) &= (\mathcal{P}_{(\lfloor t/T \rfloor - 1)T, x_1, \theta}\mathbf{u}_{\lfloor t/T \rfloor})(t \bmod T) \\ &\quad - x_1 + (\mathcal{P}_{0,x_1,\theta}\hat{\mathbf{u}})(j_0 T) \\ &= (\mathcal{P}_{(\lfloor t/T \rfloor - 1)T, x_{\lfloor t/T \rfloor}, \theta}\mathbf{u}_{\lfloor t/T \rfloor})(t \bmod T) = \hat{\mathbf{u}}(t).\end{aligned}$$

Hence, by induction we have $(\mathcal{P}_{0,x_1,\theta}\hat{\mathbf{u}})(t) = \hat{\mathbf{u}}(t)$ for all $t \in [0, JT]$. We conclude that, $\hat{\mathbf{u}}, \hat{x}, \theta$ is feasible for Problem (10) with objective value $\hat{\gamma}$. ■

To solve Optimization Problem (14), we again use a Picard contractive step to avoid explicit parametrization of the variables $\mathbf{u}_i$. However, this approach makes it difficult to directly enforce the linking constraints $\mathbf{u}_j(T) = x_{j+1}$. Our approach, then is to move this constraint into the objective by using a penalty function as

$$\min_{\substack{x \in X^{\otimes J}, \theta \in \Theta \\ \mathbf{u} \in C[0,T]^{\otimes J}}} \frac{1}{2N_s} \sum_{i=1}^{N_s} \|y_i - g(t_i, \mathbf{u}_{\lfloor t_i/T \rfloor}(t_i \bmod T), \theta)\|_2^2 + \lambda \sum_{j=1}^{J-1} \|\mathbf{u}_j(T) - x_{j+1}\|_2^2 \qquad (15)$$
$$\text{s.t. } \mathcal{P}_{(j-1)T, x_j, \theta}\mathbf{u}_j = \mathbf{u}_j \qquad \forall j \in \overline{1, J},$$

where $\lambda \ge 0$ is a regularization parameter.

Having formulated the constrained optimization problem, we now propose an extended version of the algorithm discussed in Subsec. 3.3 where Picard iterations are introduced in the objective function so that the gradient of this objective approximates the KKT conditions in Eqn. (12) – similar to the approach described in Eqn. (13). Specifically, we now have objective function

$$L_{\lambda,n}(x, \theta, \mathbf{u}) := \sum_{j=1}^{J-1} \lambda \|\mathcal{P}^n_{(j-1)T, x_j, \theta}\mathbf{u}_j(T) - x_{j+1}\|_2^2 \qquad (16)$$
$$+ \frac{1}{2N_s} \sum_{i=1}^{N_s} \|y_i - g(t_i, (\mathcal{P}^n_{(j-1)T, x_j, \theta}\mathbf{u}_{\lfloor t_i/T \rfloor})(t_i \bmod T), \theta)\|_2^2$$

for which we define Alg. 4.

Although this approach allows us to approximate the solution of a parameter estimation problem on an arbitrary time interval, there are some technical challenges. First, the approach requires additional variables x_j (initial conditions at time points $(j-1)T$) – increasing the dimension of the optimization problem. Second, the use of a penalty function in lieu of the constraint $\mathbf{u}_j(T) = x_{j+1}$ implies discontinuities unless λ is chosen to be very large. However, increasing the value of λ increases the gradient of the objective – resulting in the need for smaller step sizes and hence slowing the convergence rate of the algorithm. These numerical issues are analyzed in detail in Subsec. 7.1.

4 Gradient-Contractive Algorithms

In this section, we propose sufficient conditions for convergence of the gradient-contractive algorithms presented in Sec. 3. Specifically, we propose a general class of optimization problems which includes those posed in (10) and (15). Then, based on this generalized form of optimization problem, we define a generalized class of gradient-contractive algorithms. These two-step algorithms repeatedly update the optimization variables using both a gradient descent step and a contraction map (Algs. 2, 3 and 4 are special cases of this approach). Finally, in Thm. 8, we provide sufficient conditions such that each step of the gradient-contractive algorithm is a contraction – implying convergence.

To begin, suppose we are given a general optimization problem of the form

$$\min_{x \in X, \mathbf{u} \in U} L(x, \mathbf{u}) \quad \text{s.t. } \mathcal{P}(x, \mathbf{u}) = \mathbf{u}, \qquad (17)$$

where $U \subset \mathcal{F}(\mathbb{R}^{n_i}, \mathbb{R}^{n_u})$, $X \subset \mathbb{R}^{n_x}$, $\mathcal{P} : X \times U \to U$ and $L : X \times U \to \mathbb{R}$. Now the gradient-contractive algorithm for step sizes $\alpha > 0$, $\sigma \in (0, 1]$ and order $n \in \mathbb{N}$ is defined by the sequence $(x_k, \mathbf{u}_k) = \mathcal{T}^k(x_0, \mathbf{u}_0)$ where we say $(x_{k+1}, \mathbf{u}_{k+1}) = \mathcal{T}(x_k, \mathbf{u}_k)$ if

$$\begin{aligned} x_{k+1} &= \Pi_X\left[x_k - \alpha \nabla_x L(x_k, \mathcal{P}^n(x_k, \mathbf{u}_k))\right] \\ \mathbf{u}_{k+1} &= (1 - \sigma)\mathbf{u}_k + \sigma \mathcal{P}(x_{k+1}, \mathbf{u}_k). \end{aligned} \qquad (18)$$

Clearly, Algs. 3 and 4 are special cases of this generalized gradient contractive algorithm defined by the mapping $\mathcal{T}$, where $n = 0$ for Alg. 2. In this section, we provide conditions under which $\mathcal{T}$ is, itself, a contraction. Specifically, if we have a metric on $H := X \times U$, then we may use the following definition.

Definition 5 (Contraction) *Given a metric space H, we say $\mathcal{T} : H \to H$ is a contraction if there exists $q \in$*

$[0,1)$ *such that* $\|\mathcal{T}u_1 - \mathcal{T}u_2\|_H \leq q\|u_1 - u_2\|_H$ *for all* $u_1, u_2 \in H$.

The key property of contractive mappings is that, when iteratively applied to any $u \in H$, they converge to a *fixed point*, $u^* = \lim_{k\to\infty} \mathcal{T}^k u_0$, for which $\mathcal{T}u^* = u^*$.

Theorem 6 (Banach fixed-point theorem) *Let* H *be a complete metric space and* $T : H \to H$ *be a contraction. Then there exists unique fixed-point* $u^* \in H$ *such that* $\mathcal{T}u^* = u^*$. *Furthermore,* $\lim_{k\to\infty} \mathcal{T}^k u_0 = u^*$ *for any* $u_0 \in H$.

Thm. 8 gives sufficient conditions, expressed in terms of the properties of $\mathcal{P}$ and L, under which $\mathcal{T}$ is a contraction and, which, therefore, implies that $\mathcal{T}^k$ converges to a solution of Problem (17). First, we assume that L is separable

Definition 7 *Suppose* $U \subseteq \mathcal{F}(Z, \mathbb{R}^n)$. *We say* $L : U \to \mathbb{R}$ *is Lipschitz separable if there exist* $L_k : \mathbb{R}^n \to \mathbb{R}$ *and* $z_k \in Z$ *such that* $L(\mathbf{u}) = \sum_k L_k(\mathbf{u}(z_k))$ *and where* ∇L_k *exist and are locally Lipschitz.*

Theorem 8 *Let* $\mathcal{U} \subset \mathcal{F}(Z, \mathbb{R}^{n_u})$ *be a complete metric space,* $U \subset \mathcal{U}$ *and* $X \subset \mathbb{R}^{n_x}$ *be convex and compact sets. Suppose we have the following*

(1) $\mathcal{P} : X \times U \to U$ *is differentiable in* x *and* $\mathcal{P}(x, \mathbf{u})$ *is a contraction in* $\mathbf{u}$, *uniformly in* $x \in X$.

(2) $\nabla_x \mathcal{P}(x, \mathbf{u})$ *is Lipschitz continuous on* $(x, \mathbf{u}) \in X \times U$ *and there exist* $q < 1$, $K > 0$ *such that for all* $n \in \mathbb{N}$, $x_1, x_2 \in X$ *and* $\mathbf{u}_1, \mathbf{u}_2 \in U$

$$\|\nabla_x \mathcal{P}^n(x_1, \mathbf{u}_1) - \nabla_x \mathcal{P}^n(x_2, \mathbf{u}_2)\| \leq q^n \|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K\|x_1 - x_2\|_X. \quad (19)$$

(3) $L(x, \mathbf{u})$ *is Lipschitz separable (Def. 7) and for any* $m \in \mathbb{N}$ *and* $\mathbf{u} \in U$, $L(x, \mathcal{P}^m(x, \mathbf{u}))$ *is strongly convex on* X *uniformly in* $\mathbf{u}$ *and* n.

Then there exist $\alpha > 0$, $\sigma \in (0, 1]$ *and* $n \in \mathbb{N}$ *such that if* $\mathcal{T}$ *is defined as in* (18)*, we have the following*

a) $\mathcal{T} : X \times U \to X \times U$ *and is contractive.*

b) There exists $\nu < 1$ *and* $(x^*, \mathbf{u}^*) \in X \times U$ *such that for any* $(x_0, \mathbf{u}_0) \in X \times U$, $\lim_{k\to\infty} \mathcal{T}^k(x_0, \mathbf{u}_0) = (x^*, \mathbf{u}^*)$ *and*

$$\|\mathcal{T}^k(x_0, \mathbf{u}_0) - (x^*, \mathbf{u}^*)\|_{X\times\mathcal{U}} \leq \nu^k(\|x_0 - x^*\|_X + \|\mathbf{u}_0 - \mathbf{u}^*\|_{\mathcal{U}}).$$

PROOF. The proof consists of three parts. First, we show that by condition 2), $\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$ is Lipschitz continuous such that as $n \to \infty$, the dependence of the Lipschitz constant on $\mathbf{u}$ decreases as n increases. Second, we consider the iterate map $\mathcal{T}$ and give conditions on α, σ and n under which this map is contractive. Next, we show the existence of step sizes $\alpha, \sigma, n \in \mathbb{N}$ such that these parameters satisfy the conditions given in the second part of the proof. Finally, part b) of the theorem statement follows directly from contractivity of $\mathcal{T}$.

Variation of $\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$ **with u.** By 1) $\mathcal{P}(x, \mathbf{u})$ is contractive in $\mathbf{u}$, uniformly in $x \in X$ with factor $q_1 < 1$. Furthermore, since $\nabla_x \mathcal{P}(x, \mathbf{u})$ is Lipschitz continuous in x and U is compact, $\mathcal{P}(x, \mathbf{u})$ is Lipschitz continuous in x, uniformly in $\mathbf{u}$ (with factor K_1). Thus from the triangle inequality, for any $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$,

$$\|\mathcal{P}(x_1, \mathbf{u}_1) - \mathcal{P}(x_2, \mathbf{u}_2)\|_{\mathcal{U}} \leq q_1\|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K_1\|x_1 - x_2\|_X.$$

Now, using q, K as in Condition 2) define $q_{\mathcal{P}} = \max\{q_1, q\}$ and $K_{\mathcal{P}} = \max\left\{\frac{K_1}{1-q_1}, K\right\}$. Then from Eqn. (19), for all $n \in \mathbb{N}$, $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$,

$$\|\mathcal{P}^n(x_1, \mathbf{u}_1) - \mathcal{P}^n(x_2, \mathbf{u}_2)\|_{\mathcal{U}} \leq q_{\mathcal{P}}^n\|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K_{\mathcal{P}}\|x_1 - x_2\|_X$$

$$\|\nabla_x \mathcal{P}^n(x_1, \mathbf{u}_1) - \nabla_x \mathcal{P}^n(x_2, \mathbf{u}_2)\|_{\mathcal{U}} \leq q_{\mathcal{P}}^n\|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K_{\mathcal{P}}\|x_1 - x_2\|_X.$$

Furthermore, note that the uniform Lipschitz bound on $\mathcal{P}(x, \mathbf{u})$ also bounds the gradient as $\|\nabla_x \mathcal{P}(x, \mathbf{u})\| \leq K_1 \leq K_{\mathcal{P}}$.

Next, we show that $\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$ is Lipschitz. Note, since $L(x, \mathbf{u})$ is Lipschitz separable, there exists $\{z_i \in Z\}_i$ and $L_i : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ such $L(x, \mathbf{u}) = \sum_i L_i(x, \mathbf{u}(z_i))$. For notational clarity, we explicitly label the Lipschitz continuous partial gradients as $\nabla_1 L_i(x, v) := \nabla_x L_i(x, v)$ and $\nabla_2 L_i(x, v) := \nabla_v L_i(x, v)$. Then by the chain rule, for any $n \in \mathbb{N}$, $x \in X$ and $\mathbf{u} \in U$ we have

$$\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u})) = \sum_i \big[\nabla_1 L_i(x, \mathcal{P}^n(x, \mathbf{u})(z_i)) + \nabla_2 L_i(x, \mathcal{P}^n(x, \mathbf{u})(z_i))\nabla_x \mathcal{P}^n(x, \mathbf{u})(z_i)\big]. \quad (20)$$

Note, since X, U are compact sets and L is Lipschitz separable, $\nabla_j L_i(x, v)$ are bounded and Lipschitz continuous functions of x, v for $j \in \{1, 2\}$. Thus, for some constant $K_2 > 0$, for all $n \in \mathbb{N}$, $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $j \in \{1, 2\}$ we have

$$\begin{aligned}&\|\nabla_j L_i(x_1, \mathcal{P}^n(x_1, \mathbf{u}_1)) - \nabla_j L_i(x_2, \mathcal{P}^n(x_2, \mathbf{u}_2))\|_2 \\ &\leq K_2(\|x_1 - x_2\| + \|\mathcal{P}^n(x_1, \mathbf{u}_1)(z_i) - \mathcal{P}^n(x_2, \mathbf{u}_2)(z_i)\|_2) \\ &\leq q_{\mathcal{P}}^n K_2\|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K_2(1 + K_{\mathcal{P}})\|x_1 - x_2\|_2.\end{aligned}$$

Therefore, from Eqn. (20) and Condition (19), this implies that $\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$ is Lipschitz continuous with respect to x and $\mathbf{u}$ (as the product and sum of Lipschitz and bounded functions) and where the Lipschitz factor with respect to $\mathbf{u}$ decreases with increasing n. Specifically, there exists $K_L \geq 0$ and $\{K_{n,\mathbf{u}}\}_{n=0}^\infty \subset \mathbb{R}_+$ such that $\lim_{n\to\infty} K_{n,\mathbf{u}} = 0$ and

$$\|\nabla_x L(x_1, \mathcal{P}^n(x_1, \mathbf{u}_1)) - \nabla_x L(x_2, \mathcal{P}^n(x_2, \mathbf{u}_2))\|_2 \leq K_{n,\mathbf{u}}\|\mathbf{u}_1 - \mathbf{u}_2\|_{\mathcal{U}} + K_L\|x_1 - x_2\|_X$$

for all $n \in \mathbb{N}$, $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$.

Sufficient Conditions for $\mathcal{T}$ **to be Contractive.** Let us denote the first mapping in Eqn. (18) $\mathcal{D}_n : X \times U \to X$, such that $\mathcal{D}_n(x, \mathbf{u}) := \Pi_X[x - \alpha\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$. This map $\mathcal{D}_n$ is Lipschtitz continuous since the convexity of X implies Lipschitz continuity of Π_X and for all n, $\nabla_x L(x, \mathcal{P}^n(x, \mathbf{u}))$ is Lipschitz as shown previously. Specifically, for all n and $\alpha \geq 0$ there exists $q_{\mathcal{D},n} \geq 0$ and $K_{\mathcal{D},n} \geq 0$ such that

$$\|\mathcal{D}_n(x_1, \mathbf{u}_1) - \mathcal{D}_n(x_2, \mathbf{u}_2)\| \leq q_{\mathcal{D},n}\|x_1 - x_2\| + K_{\mathcal{D},n}\|\mathbf{u}_1 - \mathbf{u}_2\|$$

for all $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$.

Now, the map $\mathcal{T}$ defined as in (18) can be expressed as

$$\mathcal{T}(x, \mathbf{u}) = \begin{bmatrix} \mathcal{D}_n(x, \mathbf{u}) \\ (1 - \sigma)\mathbf{u} + \sigma\mathcal{P}(\mathcal{D}_n(x, \mathbf{u}), \mathbf{u}) \end{bmatrix}.$$

Note that, since $\mathcal{P}(\mathcal{D}_n(x, \mathbf{u}), \mathbf{u}) \in U$, $\mathbf{u} \in U$ and U is convex, $\mathcal{T}$ maps $X \times U$ to itself.

Next, we show that $\mathcal{T}$ is Lipschitz continuous. For all $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in U$ we have

$$\|\mathcal{T}(x_1, \mathbf{u}_1) - \mathcal{T}(x_2, \mathbf{u}_2)\|_{X\times\mathcal{U}} \leq \|\mathcal{T}(x_1, \mathbf{u}_1) - \mathcal{T}(x_2, \mathbf{u}_1)\|_{X\times\mathcal{U}} + \|\mathcal{T}(x_2, \mathbf{u}_1) - \mathcal{T}(x_2, \mathbf{u}_2)\|_{X\times\mathcal{U}}.$$

Thus, for all $x_1, x_2 \in X$ and $\mathbf{u}_1 \in \mathcal{U}$

$$\begin{aligned}\|\mathcal{T}(x_1,\mathbf{u}_1)-\mathcal{T}(x_2,\mathbf{u}_1)\|_{X\times\mathcal{U}} &= \|\mathcal{D}_n(x_1,\mathbf{u}_1)-\mathcal{D}_n(x_2,\mathbf{u}_1)\|_X \\ &+ \|\sigma\mathcal{P}(\mathcal{D}_n(x_1,\mathbf{u}_1),\mathbf{u}_1) - \sigma\mathcal{P}(\mathcal{D}_n(x_2,\mathbf{u}_1),\mathbf{u}_1)\|_{\mathcal{U}} \\ &\le q_{\mathcal{D},n}\|x_1-x_2\|_X + \sigma K_{\mathcal{P}}\|\mathcal{D}_n(x_1,\mathbf{u}_1)-\mathcal{D}_n(x_2,\mathbf{u}_1)\|_X \\ &\le q_{\mathcal{D},n}\|x_1-x_2\|_X + \sigma q_{\mathcal{D},n}K_{\mathcal{P}}\|x_1-x_2\|_X.\end{aligned}$$

Also, for all $x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$

$$\begin{aligned}&\|\mathcal{T}(x_2,\mathbf{u}_1) - \mathcal{T}(x_2,\mathbf{u}_2)\|_{X\times\mathcal{U}} \\ &\le \|\mathcal{D}_n(x_2,\mathbf{u}_1) - \mathcal{D}_n(x_2,\mathbf{u}_2)\|_{\mathcal{U}} + (1-\sigma)\|(\mathbf{u}_1-\mathbf{u}_2)\|_{\mathcal{U}} \\ &\quad + \sigma\|\mathcal{P}(\mathcal{D}_n(x_2,\mathbf{u}_1),\mathbf{u}_1) - \mathcal{P}(\mathcal{D}_n(x_2,\mathbf{u}_2),\mathbf{u}_2)\|_{\mathcal{U}} \\ &\quad\le (K_{\mathcal{D},n} + (1-\sigma) + \sigma q_{\mathcal{P}})\|\mathbf{u}_1-\mathbf{u}_2\|_{\mathcal{U}} \\ &\qquad + \sigma K_{\mathcal{P}}\|\mathcal{D}_n(x_2,\mathbf{u}_1) - \mathcal{D}_n(x_2,\mathbf{u}_2)\|_X \\ &\quad\le (1-\sigma+\sigma q_{\mathcal{P}} + K_{\mathcal{D},n} + \sigma K_{\mathcal{P}}K_{\mathcal{D},n})\|\mathbf{u}_1-\mathbf{u}_2\|_{\mathcal{U}}.\end{aligned}$$

Thus, for all $x_1, x_2 \in X$ and $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{U}$

$$\begin{aligned}\|\mathcal{T}(x_1,\mathbf{u}_1)-\mathcal{T}(x_2,\mathbf{u}_1)\|_{X\times\mathcal{U}} &\le q_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1)\|x_1-x_2\|_X \\ &+ (1-\sigma+\sigma q_{\mathcal{P}} + K_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1))\|\mathbf{u}_1-\mathbf{u}_2\|_{\mathcal{U}}.\end{aligned}$$

Therefore, $\mathcal{T}$ is a contraction if

$$(1-\sigma)+\sigma q_{\mathcal{P}}+K_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1) < 1 \quad q_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1) < 1.$$

Existence of α, σ for Contractivity of $\mathcal{T}$. First, we recall that for fixed $\mathbf{u} \in U$, if $L(x, \mathcal{P}^n(x,\mathbf{u}))$ is strongly convex on X and locally Lipschitz, then, from [31], there exists $\alpha > 0$ such that $\mathcal{D}_n(x,\mathbf{u}) : X \to X$, as defined above, is contractive in x for all fixed $\mathbf{u} \in U$. Thus, since $\mathcal{L}(x, \mathcal{P}^n(x,\mathbf{u}))$ is strongly convex on X uniformly in $\mathbf{u}$, we have that for any $n \in \mathbb{N}$, there exists $\mu > 0$ such that for any $\mathbf{u} \in U$, $L(x, \mathcal{P}^n(x,\mathbf{u}))$ is strongly convex on X with modulus $\mu \ge 0$. Thus, there exists $\alpha > 0$ such that $\mathcal{D}_n(x,\mathbf{u})$ is a contraction with factor $q_{\mathcal{D},n} \in (0,1)$ for any fixed $\mathbf{u} \in U$.

Next, α may be upper bounded by $\frac{2}{K_L}$ and $q_{\mathcal{D}_n}$ may be lower bounded by 0.5 (since it is a Lipschitz constant). Furthermore, since $L(x, \mathcal{P}^n(x,\mathbf{u}))$ is strongly convex on X uniformly in $n \in \mathbb{N}_+$, there exists $\mu > 0$ such that for all $n \in \mathbb{N}_+$ and $\mathbf{u} \in U$, $L(x, \mathcal{P}^n(x,\mathbf{u}))$ is strongly convex on X with modulus μ. Thus, $q_{\mathcal{D},n}$ may be upper bounded by $\frac{K_L-\mu}{K_L+\mu} < 1$, where K_L is Lipschitz factor of $\nabla_x L(x, \mathcal{P}^n(x,\mathbf{u}))$.

Finally, recalling $K_{n,\mathbf{u}}$ as a Lipschitz continuity bound for $\nabla_x L(x, \mathcal{P}^n(x,\mathbf{u}))$ with respect to $\mathbf{u}$, we have $\mathcal{D}_n(x,\mathbf{u})$ is Lipschitz continuous in $\mathbf{u}$ with constant $K_{\mathcal{D},n} = \alpha K_{n,\mathbf{u}}$. Thus, since $\lim_{n\to\infty} K_{n,\mathbf{u}} = 0$, there exists $n \ge 0$ such that $K_{n,\mathbf{u}} < \frac{q_{\mathcal{D},n}(1-q_{\mathcal{P}})}{\alpha}\min\left\{\frac{1-q_{\mathcal{D},n}}{q_{\mathcal{D},n}K_p}, 1\right\}$ and $\frac{K_{\mathcal{D},n}}{q_{\mathcal{D},n}(1-q_{\mathcal{P}})} = \frac{\alpha K_{n,\mathbf{u}}}{q_{\mathcal{D},n}(1-q_{\mathcal{P}})} < \min\left\{\frac{1-q_{\mathcal{D},n}}{q_{\mathcal{D},n}K_p}, 1\right\}$.

Let $\sigma \in \left(\frac{K_{\mathcal{D},n}}{q_{\mathcal{D},n}(1-q_{\mathcal{P}})}, \min\left\{\frac{1-q_{\mathcal{D},n}}{q_{\mathcal{D},n}K_p}, 1\right\}\right)$, then

$$q_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1) < q_{\mathcal{D},n}\left(\frac{1-q_{\mathcal{D},n}}{q_{\mathcal{D},n}K_p}K_p + 1\right) = 1 \quad \text{and}$$

$$\begin{aligned}1-\sigma+\sigma q_{\mathcal{P}} + K_{\mathcal{D},n}(\sigma K_{\mathcal{P}}+1) &< 1-\sigma(1-q_{\mathcal{P}}) + \frac{K_{\mathcal{D},n}}{q_{\mathcal{D},n}} \\ &< 1 - \frac{K_{\mathcal{D},n}}{q_{\mathcal{D},n}(1-q_{\mathcal{P}})}(1-q_{\mathcal{P}}) + \frac{K_{\mathcal{D},n}}{q_{\mathcal{D},n}} = 1.\end{aligned}$$

Therefore, there exists α, σ, n such that the map $\mathcal{T}$ is a contraction on $X \times U$ with factor $\nu < 1$, where ν is the maximum of the factors in the preceding two inequalities. Thus, by the Banach fixed-point theorem (Thm. 6), $\mathcal{T}$ has a unique fixed point $x^*, \mathbf{u}^*$ where

$$\|\mathcal{T}^k(x_0,\mathbf{u}_0)-(x^*,\mathbf{u}^*)\|_{X\times\mathcal{U}} \le \nu^k(\|x_0-x^*\|+\|\mathbf{u}_0-\mathbf{u}^*\|_{\mathcal{U}}).$$

■

Remark 1. The gradient-contractive approach of Thm. 8 may be extended to other gradient-based algorithms (defined by alternative maps $\mathcal{D}_n$) as long as $\mathcal{D}_n$ is contractive in x and Lipschitz in $\mathbf{u}$.

To apply the results of Thm. 8 to the parameter estimation problem, we define the mapping $\mathcal{P}$ as a Picard iteration. In the following section, we will recall properties of the Picard iteration. Then, in Sec. 6, we will give conditions on the parameterized vector field, f, under which the proposed modified and extended gradient contractive algorithms converge.

5 Contraction and Lipschitz continuity of the Picard Operator

In Sec. 4, we proposed a generalized form of gradient-contractive algorithm (Eqn. (18)) for solving the class of optimization problems defined in Eqn. (17), which are defined in terms of a Lipschitz separable objective, $L(x,\mathbf{u})$ and an equality constraint of the form $\mathcal{P}(x,\mathbf{u}) = \mathbf{u}$. Furthermore, in Thm. 8, we provided conditions on $\mathcal{P}$ and L under which the gradient-contractive algorithm converges. In this section, we return to the original parameter estimation problem, as formulated in Optimization Problems (10) and (15), where $\mathcal{P}$ is now the Picard operator. In this case, we show that Algs. 3 and 4 are special cases of the generalized gradient-contractive algorithm and that the Picard operator, $\mathcal{P}_{t_0,x,\theta}$, satisfies the conditions of Thm. 8. Specifically, in Thm. 9 and Lem. 11 (Subsec. 5.1), we show that $\mathcal{P}_{t_0,x,\theta}$ satisfies Condition 1) of Thm. 8 and in Lem. 15 (Subsec. 5.2), we show that $\mathcal{P}_{t_0,x,\theta}$ satisfies Condition 2). Next, in Sec. 6, we give conditions under which $L_{\lambda,n}$ (as defined in Eqn. (16)) likewise satisfies the conditions of Thm. 8.

5.1 *Picard Iterations as a Contraction Map*

In this subsection, we examine the Picard operator and show it satisfies Condition 1) of Thm. 8. Specifically, recall from Definition 1 that for any t_0, x, θ and parameterized vector field, f, the Picard operator is defined as

$$(\mathcal{P}_{t_0,x,\theta}\mathbf{u})(t) := x + \int_0^t f(s+t_0, \mathbf{u}(s), \theta)ds.$$

The contractive properties of the Picard operator and its gradient are relatively well-established on a sufficiently short time interval. However, in our case, we require these contractive properties to be uniform on domains $x \in X$ and $\theta \in \Theta$ which results in a slight variation on the standard result.

Theorem 9 (Picard-Lindelöf Theorem) *Suppose $\Gamma \subset \mathbb{R}$, $X \subset \mathbb{R}^{n_x}$, $\Theta \subset \mathbb{R}^{n_\theta}$ are compact and $f \in C(\Gamma \times \mathbb{R}^{n_x} \times \Theta)$ is locally Lipschitz. Let $t_0 \in \Gamma$, $a = 2\sup_{x\in X}\|x\|_X$ and $T < \min\{\frac{1}{K_x}, \frac{a}{2M}\}$ where*

$$K_x = \sup_{x,y\in B_a,\, t\in\Gamma, \theta\in\Theta} \frac{\|f(t,x,\theta) - f(t,y,\theta)\|_2}{\|x-y\|_X}$$

$$M = \sup_{x\in B_a, t\in\Gamma,\theta\in\Theta} \|f(t,x,\theta)\|_2.$$

Define $C_a := \{\mathbf{u} \in C([0,T]) \mid \|\mathbf{u}\|_\infty \le a\}$. Then for all $x \in X$ and $\theta \in \Theta$, $\mathcal{P}_{t_0,x,\theta} : C_a \to C_a$ and there exists

$\mathbf{u}^* \in C_a$ *such that for any* $\mathbf{u}_0 \in C_a$, $\lim_{k\to\infty} \mathcal{P}^k_{t_0,x,\theta}\mathbf{u}_0 = \mathbf{u}^*$. *Furthermore,* $\|\mathcal{P}_{t_0,x,\theta}\mathbf{u} - \mathcal{P}_{t_0,x,\theta}\mathbf{v}\|_\infty \le TK_x\|\mathbf{u}-\mathbf{v}\|_\infty$ *for any* $\mathbf{u},\mathbf{v} \in C_a$

PROOF. Let $x \in X$ and $\theta \in \Theta$. We first show that $\mathcal{P}_{t_0,x,\theta} : C_a \to C_a$. Recall $(\mathcal{P}_{t_0,x,\theta}\mathbf{u})(t) := x + \int_0^t f(s + t_0, \mathbf{u}(s), \theta)ds$. Suppose $\mathbf{u} \in C_a$. Since f is continuous, by, e.g. [37], $\mathcal{P}_{t_0,x,\theta}\mathbf{u}$ is continuous. Moreover,

$$\begin{aligned}\|\mathcal{P}_{t_0,x,\theta}\mathbf{u}\|_\infty &= \sup_{t\in[0,T]} \Big\|x + \int_0^t f(s+t_0,\mathbf{u}(s),\theta)ds\Big\|_2 \\ &\le \|x\|_2 + \sup_{t\in[0,T]} \Big\|\int_0^t f(s+t_0,\mathbf{u}(s),\theta)ds\Big\|_2 \\ &\le \frac{a}{2} + T \sup_{x\in B_a,\theta\in\Theta,s\in\Gamma} \|f(s,x,\theta)\|_2 \le \frac{a}{2} + TM.\end{aligned}$$

Since $T \le \frac{a}{2M}$, then $\|\mathcal{P}_{t_0,x,\theta}\mathbf{u}\|_\infty \le a$ hence $\mathcal{P}_{t_0,x,\theta}\mathbf{u} \in C_a$.

The second part of this proof is to show the contraction property. For all $\mathbf{u},\mathbf{v} \in C_a$ we have

$$\|\mathcal{P}_{t_0,x,\theta}\mathbf{u} - \mathcal{P}_{t_0,x,\theta}\mathbf{v}\|_\infty = \sup_{t\in[0,T]} \left\| \int_0^t f(s+t_0,\mathbf{u}(s),\theta) - f(s+t_0,\mathbf{v}(s),\theta)ds \right\|_2.$$

Next, for all $\mathbf{u},\mathbf{v} \in C_a$ we have

$$\begin{aligned}&\|\mathcal{P}_{t_0,x,\theta}\mathbf{u} - \mathcal{P}_{t_0,x,\theta}\mathbf{v}\|_\infty \\ &\quad\le \sup_{t\in\Gamma}\int_0^t \|f(s+t_0,\mathbf{u}(s),\theta) - f(s+t_0,\mathbf{v}(s),\theta)\|_2 ds \\ &\quad\le TK_x \sup_{s\in\Gamma}\|\mathbf{u}(s) - \mathbf{v}(s)\|_2.\end{aligned}$$

Since $TK_x < 1$, we have $\mathcal{P}_{t_0,x,\theta}$ is a contraction map. Since C_a is a compact subset of the Banach space $C([0,T])$ it is a complete metric space. Hence by the fixed point theorem, for any $\mathbf{u}_0 \in C_a$, we have $\lim_{k\to\infty}\mathcal{P}^k_{t_0,x,\theta}\mathbf{u}_0 = \mathbf{u}^* \in C_a$. ■

Thm. 9 shows that $\mathcal{P}_{t_0,x,\theta}$ satisfies Condition 1) of Thm. 8. Before moving on to prove that Condition 2) is satisfied (in Subsec. 5.2), we state two further results which will be useful in establishing this result.

First, we note that, trivially, the solution map, ϕ, itself is a fixed point of the Picard iteration, which is formally stated as follows.

Corollary 10 *Suppose the conditions of Thm. 9 are satisfied. Then for any* $x \in X$, $\theta \in \Theta$ *and for all* $\mathbf{u} \in C_a$

$$\begin{aligned}&\sup_{t\in[0,T]} \|(\mathcal{P}_{0,x,\theta}\mathbf{u})(t) - \phi(t,x,\theta)\|_2 \\ &\quad\le TK_x \sup_{t\in[0,T]} \|\mathbf{u}(t) - \phi(t,x,\theta)\|_2,\end{aligned}$$

where ϕ *is the solution map of the ODE defined by* f.

Second, we establish a uniform Lipschitz continuity bound on $\mathcal{P}_{t_0,x,\theta}$.

Lemma 11 (Lipschitz Continuity) *Suppose* Γ, X, Θ *satisfy the conditions of Thm. 9 with* t_0, a, T, K_x, M, C_a *as defined therein. Let*

$$K_\theta = \sup_{x\in B_a,\, t\in\Gamma,\, \theta_1,\theta_2\in\Theta} \frac{\|f(t,x,\theta_1) - f(t,x,\theta_2)\|_2}{\|\theta_1-\theta_2\|_2}.$$

then for any $\mathbf{u} \in C_a$ *and* $n \in \mathbb{N}$, $\mathbf{v}(t,x,\theta) := (\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)$ *is Lipschitz continuous in* $x \in X$ *and* $\theta \in \Theta$ *with bounds on the Lipschitz constants given by* $\frac{1}{1-TK_x}$ *and* $\frac{TK_\theta}{1-TK_x}$*, respectively.*

PROOF. Define the sequence $\mathbf{v}_0 = \mathbf{u}$, $\mathbf{v}_i = \mathcal{P}^i_{t_0,x,\theta}\mathbf{u}$. Clearly $\mathbf{v}_0 = \mathbf{u} \in C_a$ satisfies the Lipschitz bound since it is not a function of x, θ. Now suppose that $\mathbf{v}_i$ satisfies the Lipschitz bound. For all $x_1, x_2 \in X$, $\theta_1,\theta_2 \in \Theta$ we have

$$\begin{aligned}&\sup_{t\in[0,T]} \|\mathbf{v}_{i+1}(t,x_1,\theta_1) - \mathbf{v}_{i+1}(t,x_2,\theta_2)\|_2 \\ &= \sup_{t\in[0,T]} \Big\|x_1 - x_2 + \int_0^t f(s,\mathbf{v}_i(t_0+s,x_1,\theta_1),\theta_1) \\ &\qquad\qquad - f(t_0+s,\mathbf{v}_i(s,x_2,\theta_2),\theta_2)ds\Big\|_2 \\ &\le TK_x \sup_{t\in[0,T]} \|\mathbf{v}_i(t,x_1,\theta_1) - \mathbf{v}_i(t,x_2,\theta_2)\| \\ &\qquad\qquad + \|x_1-x_2\|_2 + TK_\theta\|\theta_1-\theta_2\|_2.\end{aligned}$$

Since $TK_x < 1$, by induction we have

$$\begin{aligned}&\sup_{t\in[0,T]} \|\mathbf{v}_n(t,x_1,\theta_1) - \mathbf{v}_n(t,x_2,\theta_2)\|_2 \\ &\qquad\le (\|x_1-x_2\|_2 + TK_\theta\|\theta_1-\theta_2\|_2)\sum_{k=0}^{n-1}(TK_x)^k \\ &\qquad\le \frac{\|x_1-x_2\|_2 + TK_\theta\|\theta_1-\theta_2\|_2}{1-TK_x}\end{aligned}$$

for all $x_1,x_2 \in X$, $\theta_1,\theta_2 \in \Theta$ and $\mathbf{u} \in C_a$. Thus, $\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}$ is Lipschitz, uniformly in $t_0, x, \theta \in \Gamma\times X\times\Theta$, and $n \in \mathbb{N}$. ■

Using these results, we now show that Condition 2) of Thm. 8 is satisfied.

5.2 Lipschitz continuity of Gradients of Picard Iterations

In this subsection, we show that $\mathcal{P}$ satisfies Condition 2) of Thm. 8. First, we use an extended version of the result in [37] (Thm. 6.20) to show that if the parameterized vector field is sufficiently smooth, the Picard operator, $\mathcal{P}$ is differentiable in x, θ.

Lemma 12 (Rudin [37]) *Let* Γ, X *and* Θ *be compact, and* $f \in C^m(\Gamma\times\mathbb{R}^{n_x}\times\Theta)$ *for* $m \in \mathbb{N}$. *If* $\mathbf{u} \in C^m(\Gamma\times X\times\Theta)$ *and* $\mathbf{v}(t,x,\theta) := (\mathcal{P}_{t_0,x,\theta}\mathbf{u})(t)$*, then* $\mathbf{v} \in C^m(\Gamma\times X\times\Theta)$.

Iteratively applying Lem. 12, we have that $\mathcal{P}^n$ is continuously differentiable with respect to $t_0, x, \theta \in \Gamma\times X\times\Theta$. If $f \in C^1$, we may thus define the gradients of an n-th order Picard operator as follows.

$$S_n(t,x,\theta) = \Big[\nabla_x(\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)\ \ \nabla_\theta(\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)\Big]. \quad (21)$$

Next, we show that $\{S_n\}_{n=1}^\infty$ is a Cauchy sequence in $C([0,T]\times X\times\Omega)$ and hence converges to the gradient of the solution map, ϕ.

Lemma 13 *Suppose* $\Gamma \subset \mathbb{R}$, $X \subset \mathbb{R}^{n_x}$, $\Theta \subset \mathbb{R}^{n_p}$ *are compact sets and* $f \in C^1(\Gamma\times\mathbb{R}^{n_x}\times\Theta)$ *is locally Lipschitz. Let* $t_0 \in \Gamma$, $a = 2\sup_{x\in X}\|x\|_2$ *and* $T < \min\{\frac{1}{K_x}, \frac{a}{2M}\}$*, where* K_x, M *are defined as in Thm. 9. Let* K_θ *be as defined in Lem. 11 and* $\nabla_x f, \nabla_\theta f$ *be locally Lipschitz continuous. Then, for any* $\mathbf{u} \in C_a$*, if* $\{S_n\}_{n=1}^\infty$ *is as defined in Eqn. (21),* $\{S_n\}_{n=1}^\infty$ *is a uniformly bounded Cauchy sequence in* $C([0,T]\times X\times\Theta)$ *with* $\|S_n\|_\infty \le 2\max\left\{\frac{1}{1-TK_x}; \frac{TK_\theta}{1-TK_x}\right\}$.

PROOF. First, we show that $S_n \in C(\Gamma\times X\times\Theta)$. Define the function $\mathbf{v}_n(t,x,\theta) := (\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)$. Then, since $f \in C^1(\Gamma\times\mathbb{R}^{n_x}\times\Theta)$ and $\mathbf{u} \in C_a$, by inductive application of Lem. 12 $\mathbf{v}_n \in C^1(\Gamma \times X \times \Theta)$. Thus, $\nabla_{x,\theta}\mathbf{v}_n \in C(\Gamma \times X \times \Theta)$.

Next, we show that $\|S_n\|_\infty \leq 2\max\left\{\frac{1}{1-TK_x}; \frac{TK_\theta}{1-TK_x}\right\}$. Since $\mathcal{P}^n_{t_0,x,\theta}\mathbf{u} \in C^1(\Gamma \times X \times \Theta)$, the Lipschitz bound in Lem. 11 implies $\left\|\nabla_{x,\theta}(\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)\right\|_2 \leq \max\left\{\frac{1}{1-TK_x}; \frac{TK_\theta}{1-TK_x}\right\}$ for all $(t,x,\theta) \in [0,T] \times X \times \Theta$.

Finally, we show that $\{S_n\}_{n=1}^\infty$ is Cauchy. From Lem. 17 (See Appendix A) there exists $K_s, M_s \geq 0$ such that for any $n, m \in \mathbb{N}$, $n \geq m$, $(x,\theta) \in X \times \Theta$ and $\mathbf{u} \in C_a$ we have

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u} - \nabla_{x,\theta}\mathcal{P}^m_{t_0,x,\theta}\mathbf{u}\|_\infty \\ \leq (TK_x)^m\left[mK_s\|\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u} - \mathbf{u}\|_\infty + \|\nabla_{x,\theta}\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u}\|_\infty\right].$$

Furthermore, $\mathbf{u} \in C_a$ and Thm. 9 imply $\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u} \in C_a$ and there exists $M_s > 0$ such that $\|\nabla_{x,\theta}\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u}\|_\infty \leq M_s$. Hence, we have

$$\sup_{t\in[0,T]}\|\nabla_{x,\theta}(\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t) - \nabla_{x,\theta}(\mathcal{P}^m_{t_0,x,\theta}\mathbf{u})(t)\|_2 \\ \leq (TK_x)^m\left[mK_{su}2a + M_s\right].$$

Since this bound holds uniformly on $x, \theta \in X \times \Theta$ we have that $\{S_n\}_{n=1}^\infty$ is Cauchy.■

An immediate consequence of Lem. 13 is that $\{S_n\}_{n=1}^\infty$ converges to the gradient of the solution map.

Proposition 14 *Suppose that the conditions of Lem. 13 are satisfied and ϕ is the solution map of the ODE defined by f. Then for all $\mathbf{u} \in C_a$ and $(t,x,\theta) \in [0,T] \times X \times \Theta$, we have $\lim_{n\to\infty}\nabla_{x,\theta}(\mathcal{P}^n_{0,x,\theta}\mathbf{u})(t) = \nabla_{x,\theta}\phi(t,x,\theta)$.*

PROOF. Since for all $\mathbf{u} \in C_a$ iterations S_n are a Cauchy sequence, then $S_n(t,x,\theta)$ converges to some $S(t,x,\theta)$ uniformly. By Thm. 7.17 in [37] we have that for all $t \in [0,T]$, $x \in X$, $\theta \in \Theta$ and for all $\mathbf{u} \in C_a$

$$\lim_{n\to\infty}\nabla_{x,\theta}(\mathcal{P}^n_{0,x,\theta}\mathbf{u})(t) = \nabla_{x,\theta}\lim_{n\to\infty}(\mathcal{P}^n_{0,x,\theta}\mathbf{u})(t) \\ = \nabla_{x,\theta}\phi(t,x,\theta),$$

where $\phi(t,x,\theta)$ is the solution map. ■

Finally, we show that $\nabla_{x,\theta}\mathcal{P}_{t_0,x,\theta}$ is Lipschitz continuous, thereby satisfying Condition 2) of Thm. 8 (wherein through some abuse of notation, $x \in X$ becomes $(x,\theta) \in X \times \Theta$ and U becomes C_a).

Lemma 15 *Suppose that the conditions of Lem. 13 are satisified. Then $\nabla_{x,\theta}\mathcal{P}_{t_0,x,\theta}\mathbf{u}$ is Lipschitz continuous on $(x,\theta) \in X \times \Theta$ and $\mathbf{u} \in C_a$ and there exist $q < 1$, $K > 0$ and $N \in \mathbb{N}$ such that for $n \geq N$, $(x_1,\theta_1), (x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u}_1, \mathbf{u}_2 \in C_a$*

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x_1,\theta_1}\mathbf{u}_1 - \nabla_{x,\theta}\mathcal{P}^n_{t_0,x_2,\theta_2}\mathbf{u}_2\|_\infty \\ \leq q^n\|\mathbf{u}_1 - \mathbf{u}_2\|_\infty + K\|x_1 - x_2\|_X + K\|\theta_1 - \theta_2\|_\Theta.$$

PROOF. Lem. 17 in Appendix A implies that, there exists $K_s \geq 0$ such that for any $n \in \mathbb{N}$, $(x,\theta) \in X \times \Theta$ and $\mathbf{u}_1, \mathbf{u}_2 \in C_a$ we have

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}_1 - \nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}_2\|_\infty \leq n(TK_x)^nK_s\|\mathbf{u}_1 - \mathbf{u}_2\|_\infty.$$

Now, let $q = \frac{TK_x+1}{2}$ and $N \geq \frac{ln(K_s)}{ln(q/TK_x)-1}$. Then, $TK_x < q < 1$ and through a tedious series of algebraic manipulations, it is straightforward to show that $n > N$ implies $n(TK_x)^nK_s < q^n$ and hence

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}_1 - \nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}_2\|_\infty \leq q^n\|\mathbf{u}_1 - \mathbf{u}_2\|_\infty,$$

which holds uniformly on $(x,\theta) \in X \times \Theta$ and $\mathbf{u}_1, \mathbf{u}_2 \in C_a$.

Next, from Lem. 18 in Appendix A, for all (x_1,θ_1), $(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$ we have

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x_1,\theta_1}\mathbf{u} - \nabla_{x,\theta}\mathcal{P}^n_{t_0,x_2,\theta_2}\mathbf{u}\|_\infty \\ \leq K\|x_1 - x_2\| + K\|\theta_2 - \theta_1\|_2.$$

We conclude that for $n > N$,

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x_1,\theta_1}\mathbf{u}_1 - \nabla_{x,\theta}\mathcal{P}^n_{t_0,x_2,\theta_2}\mathbf{u}_2\|_\infty \\ \leq q^n\|\mathbf{u}_1 - \mathbf{u}_2\|_\infty + K\|x_1 - x_2\|_2 + K\|\theta_1 - \theta_2\|_2. ■$$

To conclude, we have shown that $\mathcal{P}_{t_0,x,\theta}$ satisfies conditions 1) and 2) of Thm. 8. Next, in Sec. 6, we give conditions under which $L_{\lambda,n}$ (as defined in Eqn. (16)) likewise satisfies the conditions of Thm. 8 and use Thm. 8 to show that Algs. 3 and 4 converge.

6 Conditions for Convergence of Gradient-Contractive Algorithm

In this section, we give conditions under which Algs. 3 and 4 converge to a fixed point. Specifically, we consider Algs. 3 to be a special case of Alg. 4 and use the results of Sec. 5 to propose conditions on the loss function, $L_{\lambda,n}$ (as defined in (16)), under which the conditions of Thm. 8 are satisfied.

First, recall that the gradient-contractive algorithm for step sizes $\alpha > 0$, $\sigma \in (0,1]$ and order $n \in \mathbb{N}$ is defined by the sequence $(x_k, \mathbf{u}_k) = \mathcal{T}^k(x_0, \mathbf{u}_0)$ where we say $(x_{k+1}, \mathbf{u}_{k+1}) = \mathcal{T}(x_k, \mathbf{u}_k)$ if

$$x_{k+1} = \Pi_X\left[x_k - \alpha\nabla_x L(x_k, \mathcal{P}^n(x_k, \mathbf{u}_k))\right]$$
$$\mathbf{u}_{k+1} = (1-\sigma)\mathbf{u}_k + \sigma\mathcal{P}(x_{k+1}, \mathbf{u}_k).$$

Note that Alg. 4 is a form of gradient-contractive algorithm, with step sizes $\alpha > 0$, $\sigma \in (0,1]$ and order $n \in \mathbb{N}$, wherein $x \in X$ becomes $(x,\theta) \in X^{\otimes J} \times \Theta$ and $\mathbf{u} \in U$ becomes $\{\mathbf{u}_j\}_{j=1}^J \in C[0,T]^{\otimes J}$ and where we say $([x_{k+1}, \theta_{k+1}], \mathbf{u}_{k+1}) = \hat{\mathcal{T}}([x_k, \theta_k], \mathbf{u}_k)$ if

$$\begin{aligned} x_{k+1,j} &= \Pi_X[x_{k,j} - \alpha\nabla_{x_{k,j}}L_{\lambda,m}(x_k, \theta_k, \mathbf{u}_k)] \\ \theta_{k+1} &= \Pi_\Theta[\theta_k - \alpha\nabla_\theta L_{\lambda,m}(x_k, \theta_k, \mathbf{u}_k)] \\ \mathbf{u}_{k+1,j} &= (1-\sigma)\mathbf{u}_{k,j} + \mathcal{P}_{(j-1)T, x_{k+1,j}, \theta_{k+1}}\mathbf{u}_{k,j}. \end{aligned} \tag{22}$$

We now show that if f, g are smooth and $L_{\lambda,n}$ are strongly convex, Alg. 4 converges to a fixed point for some σ, α.

Theorem 16 *Suppose $\Gamma \subset \mathbb{R}_+$, $X \subset \mathbb{R}^{n_x}$, $\Theta \subset \mathbb{R}^{n_p}$ are compact and convex sets. Let $a = 2\sup_{x\in X}\|x\|_2$ and $T < \min\{\frac{1}{K_x}, \frac{a}{2M}\}$, where K_x, M are defined as in Thm. 9. Let $J \in \mathbb{N}$ be such that $\Gamma \subseteq [0, JT]$ and for some $\lambda > 0$, let $L_{\lambda,m}(x,\theta,\mathbf{u})$ (as defined in (16)) be strongly convex on X and Θ, uniformly in $\mathbf{u} \in C_a^{\otimes J}$ and $m \in \mathbb{N}$. Furthermore, suppose $f, g \in C^1(\Gamma \times \mathbb{R}^{n_x} \times \Theta)$ be such that $\nabla_x f, \nabla_\theta f, \nabla_x g, \nabla_\theta g$ are Lipschitz continuous.*

Then there exists $\alpha > 0$, $\sigma \in (0,1]$, $m \in \mathbb{N}$, $\nu < 1$ and $([x^*, \theta^*], \mathbf{u}^*) \in X^{\otimes J} \times \Theta \times C_a^{\otimes J}$ such that for any $(x_0, \theta_0) \in X^{\otimes J} \times \Theta$ and $\mathbf{u}_0 \in C_a^{\otimes J}$, if $([x_k, \theta_k], \mathbf{u}_k) = \hat{\mathcal{T}}^k([x_0, \theta_0], \mathbf{u}_0)$ where $\hat{\mathcal{T}}$ is as defined in Eqn. (22), we have $\lim_{k\to\infty}([x_k, \theta_k], \mathbf{u}_k) = ([x^*, \theta^*], \mathbf{u}^*)$ and

$$\|x_k - x^*\|_2 + \|\theta_k - \theta^*\|_2 + \|\mathbf{u}_k - \mathbf{u}^*\|_\infty \\ \leq \nu^k(\|x_0 - x^*\|_2 + \|\theta_0 - \theta^*\|_2 + \|\mathbf{u}_0 - \mathbf{u}^*\|_\infty).$$

PROOF. The proof uses the results of the previous section to verify the conditions of Thm. 8. First, we define $\hat{U} := C_a^{\otimes J}$ and $\hat{X} = X^{\otimes J}$ (with norms $\|\hat{x}\|_{\hat{X}} = \max_i \|\hat{x}_i\|_2$ and $\|\hat{\mathbf{u}}\|_{\hat{U}} = \max_i \|\hat{\mathbf{u}}_i\|_\infty$, respectively). Then, since X is compact, $\hat{X}$ is compact and, furthermore, $2\sup_{\hat{x}\in\hat{X}} \|\hat{x}\|_{\hat{X}} = 2\max_{j\in\overline{1,J}} \sup_{\hat{x}_j\in X} \|\hat{x}_j\|_X = a$.

Define $\hat{f} \in C([0,T]\times\mathbb{R}^{n_x J}\times\Theta)$ as $\hat{f}(t,\hat{x},\theta) = \left[\hat{f}_1^T(t,\hat{x},\theta)\ \cdots\ \hat{f}_J^T(t,\hat{x},\theta)\right]^T$, where $\hat{f}_j(t,\hat{x},\theta) = f(t+(j-1)T,\hat{x}_j,\theta)$. Then

$$\sup_{\substack{\hat{x}\in B_a,\theta\in\Theta\\ t\in[0,T]}} \|\hat{f}(t,\hat{x},\theta)\|_{\hat{X}} = \max_{j\in\overline{1,J}} \sup_{\substack{\hat{x}_j\in X,\theta\in\Theta\\ t\in[0,T]}} \|\hat{f}_j(t,\hat{x},\theta)\|_2 = M$$

$$\sup_{\substack{\hat{x},\hat{y}\in B_a,\theta\in\Theta\\ t\in[0,T]}} \frac{\|\hat{f}(t,\hat{x},\theta)-\hat{f}(t,\hat{y},\theta)\|_{\hat{X}}}{\|\hat{x}-\hat{y}\|_{\hat{X}}}$$
$$= \sup_{\substack{\hat{x},\hat{y}\in B_a^{\otimes J}\\ t\in[0,T],\theta\in\Theta}} \frac{\max_{j\in\overline{1,J}} \|\hat{f}_j(t,\hat{x}_j,\theta)-\hat{f}_j(t,\hat{y}_j,\theta)\|_X}{\max_{j\in\overline{1,J}} \|\hat{x}_j-\hat{y}_j\|_X}$$
$$\le \sup_{\substack{\hat{x},\hat{y}\in B_a^{\otimes J}\\ t\in[0,T],\theta\in\Theta}} \frac{\max_{j\in\overline{1,J}} K_x\|\hat{x}_j-\hat{y}_j\|_X}{\max_{j\in\overline{1,J}} \|\hat{x}_j-\hat{y}_j\|_X} = K_x.$$

We conclude that, $\hat{f}$, $\hat{\Gamma} = [0,T]$, $\hat{X}$, and Θ satisfy the conditions of Thm. 9 with $t_0 = 0$ and a, T, K_x, M, C_a as defined therein. We conclude that the Picard operator, $\hat{\mathcal{P}}_{0,\hat{x},\theta}$, defined by the parameterized vector field $\hat{f}$, satisfies $\|\hat{\mathcal{P}}_{0,x,\theta}\mathbf{u} - \hat{\mathcal{P}}_{0,x,\theta}\mathbf{v}\|_\infty \le TK_x\|\mathbf{u}-\mathbf{v}\|_\infty$ for all $(x,\theta)\in\hat{X}\times\Theta$ and for any $\mathbf{u},\mathbf{v}\in C_a$. This implies Condition 1) of Thm. 8 is satisfied, where Lem. 12 establishes differentiability of $\hat{\mathcal{P}}_{0,\hat{x},\theta}$.

Second, since $\hat{f}_j \in C^1(\Gamma\times\mathbb{R}^{n_x}\times\Theta)$ and $\nabla_{x,\theta}\hat{f}_j$ are Lipschitz continuous, we have that $\hat{f}$ satisfies the conditions of Lem. 13. Thus, $\nabla_{x,\theta}\hat{\mathcal{P}}_{0,x,\theta}\mathbf{u}$ is Lipschitz continuous with respect to x,θ and, from Lem. 15 there exists $q < 1, K > 0$, $N \in \mathbb{N}$ such that for $n \ge N$, $(x_1,\theta_1),(x_2,\theta_2)\in\hat{X}\times\Theta$ and $\mathbf{u}_1,\mathbf{u}_2\in\hat{U}$

$$\|\nabla_{x,\theta}\hat{\mathcal{P}}^n_{0,x_1,\theta_1}\mathbf{u}_1 - \nabla_{x,\theta}\hat{\mathcal{P}}^n_{0,x_2,\theta_2}\mathbf{u}_2\|_\infty$$
$$\le q^n\|\mathbf{u}_1-\mathbf{u}_2\|_\infty + K\|x_1-x_2\|_{\hat{X}} + K\|\theta_1-\theta_2\|_2.$$

Thus, $\nabla_{x,\theta}\hat{\mathcal{P}}_{0,x,\theta}\mathbf{u}$ satisfies Condition 2) of Thm. 8.

Finally, for given $\lambda>0$, if we define $\hat{t}_i := t_i \bmod T$ and

$$L(x,\theta,\mathbf{u}) := \frac{1}{2N_s}\sum_{i=1}^{N_s}\|y_i - g(t_i,\mathbf{u}_{\lfloor t_i/T\rfloor}(\hat{t}_i),\theta)\|_2^2 + \lambda\sum_{j=1}^{J-1}\|\mathbf{u}_j(T)-x_{j+1}\|_X^2.$$

Then $L(x,\theta,\mathbf{u}) = L_{\lambda,0}(x,\theta,\mathbf{u})$ and $L(x,\theta,\hat{\mathcal{P}}^m_{0,x,\theta}\mathbf{u}) = L_{\lambda,m}(x,\theta,\mathbf{u})$, with $L_{\lambda,m}$ as defined in (16). Since $g\in C^1(\Gamma\times\mathbb{R}^{n_x}\times\Theta)$ and $\nabla_{x,\theta}g$ is Lipschitz continuous, we conclude L is Lipschitz separable. By assumption, $L_{\lambda,m}$ is strongly convex on X and Θ uniformly in $\mathbf{u}\in\hat{U}$ and $m\in\mathbb{N}$, and hence Condition 3) of Thm. 8 is satisfied.

We conclude that convexity of X, C_a imply convexity of $\hat{X}\times\Theta$, $\hat{U}$ and hence all conditions of Thm. 8 are satisfied where $C_a \mapsto \hat{U}$ and $X \mapsto \hat{X}\times\Theta$ are convex and compact, $\mathcal{P}(x,\mathbf{u}) \mapsto \hat{\mathcal{P}}_{0,\hat{x},\theta}\mathbf{u}$, and $\mathcal{T}(x,\mathbf{u}) \mapsto \hat{\mathcal{T}}([x,\theta],\mathbf{u})$ – which concludes the proof.■

Remark 1. Note that in practice we have four hyperparameters α,σ,T and n. One option to choose these parameters is to fix $\sigma = \min\{1, \frac{\mu}{K_L-\mu}\}$ and $\alpha = \frac{2}{\mu+K_L}$, where μ is a strong convexity modulus of $L_{\lambda,n}$ and K_L is Lipschitz constant of $\nabla_{x,\theta}L_{\lambda,n}$. Then, for a fixed T, n may be chosen sufficiently large. Alternately, for a fixed $n \ge 1$ we may choose T sufficiently small.

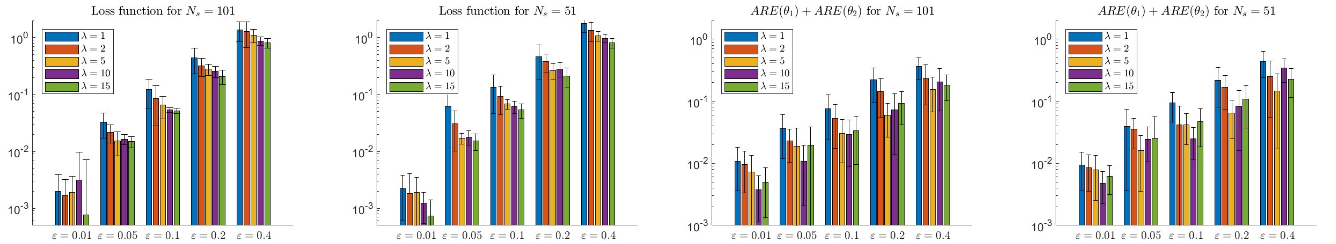
7 Numerical experiments.

Having defined a gradient contractive algorithm based on the Picard mapping (Alg. 4), and having established conditions for convergence of this algorithm to a solution of the parameter estimation problem, we now proceed to evaluate practical aspects of the performance of this algorithm. This analysis consists of five tests, each focusing on one metric of performance of the algorithm, and each using a nonlinear model and dataset indicative of that metric of performance. In all examples, initial conditions are unknown and measurements include varying levels of noise. All tests of Alg. 4 use step size $\sigma = .5$ and a backtracking line search to determine step size α. The parameter T varies by example and termination occurs when the norm of gradient $\|\nabla_\theta L_{\lambda,n}\|$ is sufficiently small. The numerical tests are summarized as follows.

First, in Subsec. 7.1, we evaluate the effect of the regularization parameter λ on error in the parameter estimates using various levels of noise and as applied to the Van der Pol Oscillator, with direct measurement of both states. Second, in Subsec. 7.2, we evaluate the effect of irregular sampling intervals on error in parameter estimates and as illustrated by the FitzHugh-Nagumo neuron using measurements of a single state. Third, in Subsec. 7.3,we consider performance when data is obtained using multiple instances but with sparse data (long periods between samples) as illustrated by the Rozenzweig MacArthur model. Fourth, in Subsec. 7.4 we consider experiments driven by external excitation and with a large number of states and parameters. We evaluate the effect of different excitations on estimation accuracy as illustrated by a model of tumor growth. Finally, in Subsec. 7.5, we compare the accuracy of the algorithm with both comparable heuristics [9] and gradient-free optimization [29, 38] as illustrated using the Lorenz model.

7.1 The Van der Pol Oscillator and the Effect of Regularization Parameter

In the first numerical example, we investigate the performance of Alg. 4 (defined by regularized loss function $L_{\lambda,n}$ in Eqn. (16)) as a function of regularization parameter λ. Recall, this regularization parameter λ weights a penalty for discontinuity of the approximated solution at times jT. Hence, for a larger weight λ, the estimated solution, $\mathbf{u}$ is closer to the solution map on extended time intervals and hence better captures the relationship between estimated parameter values and evolution of the state. However, as λ increases, convergence time



(a) Loss for $N_s = 101$. (b) Loss for $N_s = 51$. (c)Parameters Errs.$N_s = 101$. (d)Parameters Errs.$N_s = 51$.

Fig. 1. Loss function as in Eqn. (6) and normalized error of parameters (ARE) for Van der Pol Oscillator in Eqn. (23) identified by Alg. 4 for $n = 1$, $T = 0.25$ and $\lambda \in \{1, 2, 5, 10, 15\}$. The true parameters, $\theta_1 = \theta_2 = 1$, variance of the measurements is $\varepsilon \in \{0.01, 0.05, 0.1, 0.2, 0.4\}$. The number of measurements is $N_s = 101$ or $N_s = 51$. The data for $N_s = 51$ is a subsample of the measurements $N_s = 101$ for twice larger sampling times. Average and standard deviation of ARE of parameters and loss function is based on 10 trials with randomized initial conditions.

increases and less weight is placed on matching the estimated solution to the given data. To provide guidance on choice of regularization parameter, we consider Van der Pol oscillator.

$$\begin{aligned} \dot{x}_1(t) &= \theta_1 x_2(t) \\ \dot{x}_2(t) &= -\theta_1 x_1(t) + \theta_2(1 - x_1^2(t))x_2(t), \end{aligned} \tag{23}$$

where both states are observable and our goal is to estimate the parameters θ_1, θ_2. The data sets used in this example are generated from a single instance (with single initial condition x_0) over a time interval $[0, t_f]$ using $t_f = 10$ and true parameter values $\theta_1 = \theta_2 = 1$, N_s evenly distributed sampling times $t \in \{0 : t_f/(N_s - 1)i : t_f\}$ and noisy data of the form $y_i = (1 + n_i)\phi(t_i, x_0, \theta) + m_i$, where $\phi(t_i, x, \theta)$ are the actual states and n_i, m_i are normally distributed with zero mean and variance ε.

To apply Alg. 4, we use the measurement data to approximate a bound on the Lipschitz constant as $K_x \sim 4$ which yields a convergence interval of $T = 1/K_x = .25s$. This yields $J = t_f/T = 40$ sub-intervals. A single Picard iteration $n = 1$ is used.

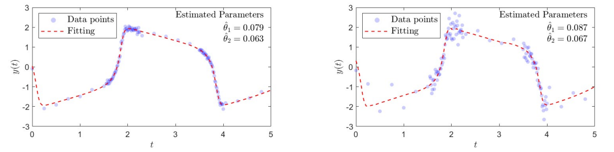
To evaluate the effect of regularization, we now construct 100 data sets, Y_{1,ε,x_0} (using $N_s = 101$) and Y_{2,ε,x_0} (using $N_s = 51$) for variances $\varepsilon \in \{0.01, 0.05, 0.1, 0.2, 0.4\}$ and 10 randomly selected initial conditions $x_0 \in X_{\lambda,\varepsilon}$ (such that x_0 is always inside the limit cycle). Alg. 4 is then applied to each data set using regularization parameters $\lambda \in \{1, 2, 5, 10, 15\}$ – yielding parameter estimates $\hat{\theta}_{\lambda,\varepsilon,x_0}$ and initial condition estimates $\hat{x}_{\lambda,\varepsilon,x_0}$.

For each choice of λ, ε, performance is evaluated by first performing numerical simulation with estimated parameters, $\hat{\theta}_{\lambda,\varepsilon,x_0}$, and initial conditions, $\hat{x}_{\lambda,\varepsilon,x_0}$, and then computing the average loss so that if ϕ is the solution map, we have

$$L(\lambda, \varepsilon) = \frac{1}{10} \sum_{x_0 \in X_{\lambda,\varepsilon}} \frac{1}{N_s} \sum_{i=1}^{N_s} \|\phi(t_i, \hat{x}_{\lambda,\varepsilon,x_0}, \hat{\theta}_{\lambda,\varepsilon,x_0}) - y_i\|_2^2.$$

In addition, for a parameter, θ, we compute the average relative error (ARE) in the parameter estimates as

$$ARE(\lambda, \varepsilon) = \frac{1}{10} \sum_{x_0 \in X_{\lambda,\varepsilon}} \frac{|\theta - \hat{\theta}_{\lambda,\varepsilon,x_0}|}{|\hat{\theta}_{\lambda,\varepsilon,x_0}|}. \tag{24}$$



(a)Results for variance ε=.05. (b)Results for variance ε=.2.

Fig. 2. Parameter Estimation using Alg. 4 for the FitzHugh–Nagumo model in Eqn. (25) with irregular sampling times (blue dots) clustered around spiking and bursting with noise variances of $\varepsilon = .05$ (Fig 2a) and $\varepsilon = .2$ (Fig 2b). The red dashed line indicates simulation based on the estimated parameters ($\hat{\theta}_i$) and initial states. True parameter values are $\theta_1 = 0.08$ and $\theta_2 = 0.064$.

As seen in Fig. 1, increasing the regularization parameter λ significantly improves both average loss and ARE for large noise variance ($\varepsilon \in \{0.1, 0.2, 0.4\}$), but does not significantly improve performance with low noise ($\varepsilon \in \{0.01, 0.05\}$). In addition, as expected, more data ($N_s = 101$ vs $N_s = 51$) improved performance in almost every case – See Fig. 1 a, c) and b, d).

7.2 *The FitzHugh-Nagumo Neuron and the Effect of Irregular Sampling*

In this subsection, we evaluate the effect of irregular sampling intervals on error in estimated parameters. For this case, we consider the FitzHugh-Nagumo model a single cortical neuron, which represents spiking and bursting behavior of neurons [39]. This model is a standard example of a stiff ODE and hence it is known that sampling at times of spiking and bursting are more useful than during interim periods – See Fig. 2. Specifically, we consider the model

$$\begin{aligned} \dot{x}_1(t) &= 10(x_1(t) - 1/3x_1^3(t) - x_2(t) + I(t)) \\ \dot{x}_2(t) &= 10(\theta_1 x_1(t) - \theta_2 x_2(t) + 0.056) \\ y(t) &= g(x(t)) := x_1(t), \end{aligned} \tag{25}$$

where the states, $x_1(t), x_2(t) \in \mathbb{R}$, represent membrane potential of neurons and membrane recovery variables and the input, $I(t)$, represents the stimulus current and is set to $I(t) = 1$ when the neuron is active. The measured output is $y(t)$ and the parameters to be estimated are $\theta_1, \theta_2 \geq 0$.

For this test, we use true parameter values $\theta_1 = 0.08$ and $\theta_2 = 0.064$. Based on observations of the output, $y(t) = x_1(t)$, sampling times are clustered around observed spiking and bursting times, $T_2 := \{1.55 + 0.0125k\}_{k=0}^{56}$ and $T_4 := \{3.5 + 0.0125k\}_{k=0}^{40}$. Sampling times during recovery periods are less frequent: $T_1 := \{0.25k\}_{k=0}^{6}$, $T_3 := \{2.3 + 0.25k\}_{k=0}^{4}$, and $T_5 := \{4.05 + 0.25k\}_{k=0}^{3}$ where the union of all sampling times is defined as $t_i \in T := \cup_{i=1:5} T_i$. These sampling times are depicted as blue dots in Fig. 2 and where (as in Subsec. 7.1) measurements are taken over a single instance using initial condition $x_1(0) = 0, x_2(0) = 2$ as

$$y_i = (1 + n_i)g(\phi(t_i, x(0), \theta)) + m_i \qquad t_i \in T,$$

where ϕ is the solution map and n_i, m_i are normally distributed with zero mean and variances of $\varepsilon = .05$ (Fig 2a) and $\varepsilon = .2$ (Fig 2b). Applying Alg. 4 with $\lambda = 5$, $n = 2$, and $T = .2$ we obtain parameter estimates $\hat{\theta}_1 = 0.079$ and $\hat{\theta}_2 = 0.063$ for variance $\varepsilon = .05$; and $\hat{\theta}_1 = 0.087$ and $\hat{\theta}_2 = 0.067$ for variance $\varepsilon = .2$.

7.3 Rosenzweig-MacArthur predator-prey model and the Effect of Sparse Data.

In the next example, we evaluate performance of Alg. 4 for sparse data using the Rosenzweig-MacArthur predator-prey model [40], wherein the states represent populations of predators and times are measured in years. As in many biological models, collecting data entails the labor-intensive counting of populations and hence data for such models is typically limited to a few samples. In addition, measurements of biological processes are often inaccurate and hence repeated experiments are used to obtain statistically significant sample sizes. For example, in the case of predator-prey models, data may be based on tracking several distinct populations in order to reduce the influence of unmodelled and localized factors such as geography. In our case, the model takes the following form [40]

$$\dot{x}_1(t) = rx_1(t)\left(1 - \frac{x_1(t)}{K}\right) - \frac{ax_1(t)x_2(t)}{b + x_1(t)}$$
$$\dot{x}_2(t) = k\frac{ax_1(t)x_2(t)}{b + x_1(t)} - cx_2(t), \tag{26}$$

where $x_1(t), x_2(t)$ represent prey and predator populations, respectively. This model has 6 unknown parameters (a, b, c, r, k, K), with: r, K representing reproduction rate and carrying capacity of the prey population; a, k are encounter and growth rates, b is an environmental factor, and c is predator death rate. For this test, we use true parameters $a = 2.8$, $b = 0.7$, $c = 1.35$, $r = 3.5$, $k = 1.5$ and $K = 1.4$.

To evaluate the effect of sparse sampling and multiple instances, the algorithm is tested using data obtained from between 1 and 4 distinct populations ($N_c = 1:4$ is number of initial conditions), with population counts being made at years $t_i \in \{0, 1.25, 2.5, 3.75, 5\}$ using measurement with noise model $y_{ij} = (1+n_i)\phi(t_i, x_j, \theta)+m_i$, where n_i, m_i are normally distributed with zero mean and variance $\varepsilon = 0.05$.

Alg. 4 is then applied to each case with parameters $\lambda = 5$, $n = 2$ and $T = 0.2$. The test is repeated ten

Identified Parameters	Available Inputs u_1	$u_{1:2}$	$u_{1:3}$	$u_{1:4}$	$u_{1:5}$
$ARE(\theta_1)$	0.34 ± 0.06	0.18 ± 0.05	0.23 ± 0.15	0.12 ± 0.02	0.03 ± 0.03
$ARE(\theta_2)$	0.44 ± 0.09	0.21 ± 0.07	0.43 ± 0.13	0.28 ± 0.04	0.07 ± 0.07
$ARE(\theta_3)$	0.67 ± 0.16	0.31 ± 0.13	0.88 ± 0.05	0.67 ± 0.08	0.15 ± 0.16
$ARE(\theta_4)$	0.43 ± 0.09	0.20 ± 0.07	0.33 ± 0.18	0.17 ± 0.04	0.05 ± 0.03
$ARE(\theta_5)$	0.78 ± 0.32	2.70 ± 1.50	3.96 ± 0.00	3.96 ± 0.00	1.39 ± 0.99
$ARE(\theta_6)$	0.70 ± 0.18	0.37 ± 0.03	0.78 ± 0.18	0.57 ± 0.08	0.01 ± 0.01
$ARE(\theta_7)$	0.66 ± 0.00	0.69 ± 0.02	0.99 ± 0.00	0.99 ± 0.00	0.96 ± 0.05

Table 2
Normalized error in identified parameters $(\theta_1, ..., \theta_6, \rho)$ for model of tumor growth as in (27) based on multiple input-output measurements. Each pairs of input-output measurement has $N_s = 51$ data samples and is generated by input u from the set of observed inputs $u_{1:5}$ with noise variance 0.01. True parameters are $\theta_1 = 1$, $\theta_2 = 1.5$, $\theta_3 = 0.5$. $\theta_4 = 0.33$, $\theta_5 = 0.01$, $\theta_6 = 0.2$ and $\rho = 0.3$. The initial condition $N(0) = 1$, $T(0) = 2$, $I(0) = 1.65$, $v(0) = 0$ is identical for each input-output pair. Average and standard deviation of parameter errors are presented based on 3 different trials. The green values represent that the parameter has been identified ($ARE < 0.3$) and the red values represent that the parameter has not been identified ($ARE > 0.3$).

times and the average relative error in the parameter estimates is listed in Table 1, where even for small amount of samples we were able to identify parameters. As expected, these results show that increasing the number of populations sampled reduces error in the parameter estimates, with reasonable accuracy obtained using as few as 2 populations.

7.4 Tumor Growth, Identifiability, and Excitation.

In this example, we examine the significance of model excitation for a weakly identifiable model. As discussed in [41], biological models with multiple states and parameters are often weakly identifiable in that multiple choices of parameters yield the same or approximately the same solution. We examine here the effect of using datasets generated from multiple non-redundant excitations to improve identifiability of these parameters. Specifically, we use a model of tumour immunodynamics with chemotherapy [42] where different drug dosing strategies are used as inputs to the model, given as

$$\dot{N}(t) = \theta_1 N(t)(1 - N(t)) - \theta_1 N(t)T(t) - (1 - e^{-v(t)})N(t)$$
$$\dot{T}(t) = \theta_2 T(t)(1 - T(t)) - \theta_3 T(t)I(t) - \theta_1 T(t)N(t) - 3(1 - e^{-v(t)})T(t)$$
$$\dot{I}(t) = \theta_4 + \theta_5 \frac{I(t)T(t)}{\rho + T(t)} - \theta_1 I(t)T(t) - \theta_6 I(t) - 2(1 - e^{-v(t)})I(t) \tag{27}$$
$$\dot{v}(t) = -v(t) + u(t)$$
$$y(t) = g(N(t), T(t), I(t), v(t)) := \begin{bmatrix} N(t) & T(t) & I(t) \end{bmatrix}^T,$$

where $N(t)$ is the density of healthy cells, $T(t)$ is tumor cells, $I(t)$ is immune cells and $v(t)$ is the concentration of drugs in the tumor area. The system has a single input $u(t)$, that represents drug dosing. We assume that states $N(t), T(t), I(t)$ are outputs of the system. Parameters

# of Initial Conditions	Total Number of Samples	Identified Parameters $ARE(a)$	$ARE(b)$	$ARE(c)$	$ARE(r)$	$ARE(k)$	$ARE(K)$	$\sum_{i=1}^{6}\frac{ARE(\theta_i)}{6}$
$N_c = 1$	$N_sN_c = 5$	0.19±0.11	0.37±0.29	0.14±0.12	0.35±0.36	0.30±0.31	0.21±0.15	0.26±0.18
$N_c = 2$	$N_sN_c = 10$	0.18±0.19	0.17±0.11	0.10±0.09	0.13±0.09	0.17±0.15	0.07±0.06	0.14±0.09
$N_c = 3$	$N_sN_c = 15$	0.17±0.15	0.17±0.15	0.09±0.06	0.08±0.09	0.14±0.12	0.06±0.05	0.12±0.08
$N_c = 4$	$N_sN_c = 20$	0.13±0.11	0.15±0.16	0.07±0.06	0.06±0.05	0.11±0.10	0.06±0.05	0.10±0.07

Table 1
Normalized error in estimated parameters using Alg. 4 for the Rosenzweig-MacArthur predator-prey model (26) vs. number of populations sampled using sparse sampling times $t \in \{0, 1.25, 2.5, 3.75, 5\}$. Each test is repeated 10 times using randomized initial conditions noise variance $\varepsilon = .05$ and Average Relative Error (ARE) and standard deviation are listed. True parameter values are $a = 2.8$, $b = 0.7$, $c = 1.35$, $r = 3.5$, $k = 1.5$, $K = 1.4$.

values are taken from [42], where $\theta_1 = 1$, $\theta_2 = 1.5$, $\theta_3 = 0.5$. $\theta_4 = 0.33$, $\theta_5 = 0.01$, $\theta_6 = 0.2$ and $\rho = 0.3$. The initial conditions are $N(0) = 1$, $T(0) = 2$, $I(0) = 1.65$, $v(0) = 0$.

5 drug dosing strategies are chosen as $u_1(t) = 0$,

$$u_2(t) = \begin{cases} e^{-t}, & \text{if } t \in [0,3] \\ 0, & \text{otherwise} \end{cases} \quad u_3(t) = \begin{cases} e^{t}, & \text{if } t \in [0,3] \\ 0, & \text{otherwise} \end{cases}$$

$$u_4(t) = \begin{cases} 3-t & \text{if } t \in [3,6] \\ 0, & \text{otherwise} \end{cases} \quad u_5(t) = \begin{cases} e^{t}, & \text{if } t \in [1.5, 4.5] \\ 0, & \text{otherwise,} \end{cases}$$

where it is observed that u_3, u_4, u_5 clear the tumour, but u_1 and u_2 do not.

For each dosing strategy, $i = 1, \cdots, 5$, we generate 3 datasets $\{Y_{i,j}\}_{j=1}^{3}$ each using measurement noises n_i, m_i (normally distributed with zero mean and variance $\varepsilon = .01$) so that $y_i = (1 + n_i)g(\phi_u(t_i, x, \theta)) + m_i$ where ϕ_u is the solution of Eqn. (27) for input $u(t)$, and $t_i = \{0 : \frac{t_f}{N_s-1} i : t_f\}$, where $N_s = 51$ and $t_f = 6$.

For testing, we apply Alg. 4 (with $\lambda = 2$, $n = 2$ and $T = 0.1$) to the datasets $Y_{1,j}$, $Y_{1,j} \cup Y_{2,j}$, $Y_{1,j} \cup Y_{2,j} \cup Y_{3,j}$, $Y_{1,j} \cup Y_{2,j} \cup Y_{3,j} \cup Y_{4,j}$, and $Y_{1,j} \cup Y_{2,j} \cup Y_{3,j} \cup Y_{3,j} \cup Y_{5,j}$ – representing an increase in the number of excitations used to identify the data. In each case we determine the ARE (over $j = 1 : 3$) for each identified parameter. The results are listed in Table 2, where green indicates that the ARE for the given parameter is less that 0.3. These results indicate the failure to include any excitation (Y_{1j}) results in failure to effective identify the parameters while inclusion of inputs u_2, u_3, u_4 generates improved results. Inclusion of all 5 excitations generates the best results, although even in this case, it seems that some parameters are not identifiable – even though true values of the parameters are not found in this case, numerical simulation using identified parameters closely resembles the experimental data. In such cases, reparameterization of the model may be needed [41].

7.5 The Lorenz System: Chaos and Comparison.

The goal of this last numerical test is to evaluate the accuracy of estimated parameters in the proposed algorithms as compared to existing state-of-the-art methods for parameter estimation. For this analysis we use data generated from the well-studied Lorenz system [3, 43], originally proposed as a model of convection rolls in the atmosphere [44]. Solutions to the Lorentz system tend to converge to a 2D manifold known as a "strange attractor", but are chaotic within this manifold – implying adjacent trajectories diverge exponentially fast. As a result, while parameter estimation for the Lorentz system is reasonably accurate when measurement noise is small, many methods for parameter estimation fail when noise is significant. The Lorenz system is defined as

$$\begin{aligned} \dot{x}_1(t) &= 10\sigma(x_2(t) - x_1(t)) \\ \dot{x}_2(t) &= 10x_2(t)(\rho - x_3(t)) - x_2(t) \\ \dot{x}_3(t) &= 10x_1(t)x_2(t) - \beta x_3(t), \end{aligned} \tag{28}$$

where x_1, x_2, x_3 are states and σ, ρ, β are parameters.

For this analysis, we use Lorenz parameters $\sigma = 1$, $\beta = \frac{8}{3}$ and $\rho = 2.8$. In order to allow for comparison with existing methods, each dataset uses a single intial condition, and directly measures the state at frequent, regular sampling times $t_i = \{0 : 2i/(N_s - 1) : 2\}$ ($N_s = 101$). Measurement noise is modelled as $y_i =$

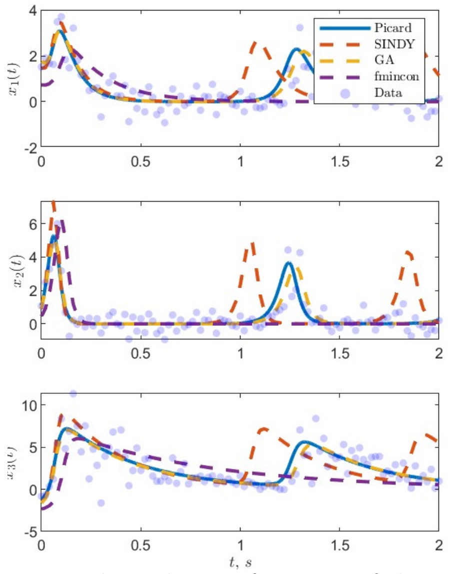
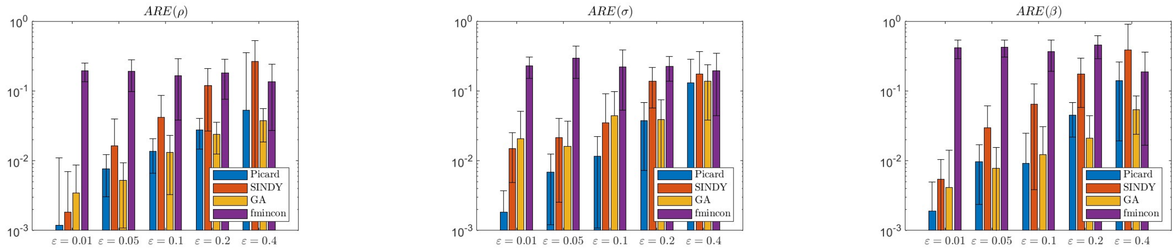


Fig. 4. Numerical simulation of 3 states of the Lorenz system (Eqn. (28)) using identified parameters from Picard (Alg. 4), SINDy, fmincon and GA alternatives (See Fig. 3 and description) for noise variance $\varepsilon = 0.4$. True parameter values are $\rho = 2.8$, $\sigma = 1$, $\beta = 8/3$; initial states are $x(0)$=$[1.46, 0.90, -1.57]^T$; and blue dots are measurements.



(a) Error in identified ρ parameter. (b) Error in identified σ parameter. (c) Error in identified β parameter.

Fig. 3. Accuracy in identified system parameters for the Lorenz system (Eqn. (28)) using Alg. 4 (Picard), as compared with: SINDy; MATLAB's `nlgreyest` (w. fmincon) (fmincon); and a black box gradient-free alternative approach to solving Eqn. (6) (GA). For each algorithm, ARE and standard deviations for each parameter ρ (Fig. 3a), σ (Fig. 3b), β (Fig. 3c) are computed over 10 datasets for each level of measurement noise variance – $\varepsilon \in \{0.01, 0.05, 0.1, 0.2, 0.4\}$. True parameter values are $\rho = 2.8$, $\sigma = 1$, $\beta = 8/3$.

$(1+n_i)\phi(t_i, x_0, \theta) + m_i$, where n_i, m_i are normally distributed with zero mean and variance ε. This approach is used to obtain 50 data sets $Y_{\varepsilon,x_0} = \{(y_i, t_i)\}_{i=1}N_s$, where $\varepsilon \in \{0, 0.01, 0.05, 0.1, 0.2, 0.4\}$ and 10 randomly chosen initial conditions x_0, normally distributed with zero mean and variance 1.

Parameter estimation is then performed on each dataset using each of 4 algorithms: Picard (Alg. 4); SINDy [3]; fmincon [21]; and GA [29]. For Picard, Alg. 4 is used with parameters $\lambda = 5$, $n = 2$ and $T = 0.2$. For SINDy, we use the a slight modification of the black-box implementation given in [3] (regularization parameter $\lambda = 0.25$ and measurements of $\dot{x}(t_i)$ are computed using first-order difference). For fmincon, we use `nlgreyest` from MATLAB's System Identification Toolbox [38] with selected optimizer: `fmincon`. For GA, we designed a gradient-free black-box optimization approach to solution of Eqn. (6). Specifically, for given x, θ, we use numerical simulation to compute the solution ϕ which then can be used to compute the least squares loss. With this loss oracle in hand, we use a genetic algorithm (MATLAB's `ga` function – with tolerances 0.01 and run time of 5 mins on Intel i7-4960X CPU at 3.60 GHz) to solve the optimization problem. For all methods, parameters are restricted to $\rho \in [1.4, 4.2]$, $\sigma \in [0.5, 1.5]$, $\beta \in [\frac{4}{3}, 4]$, and $x_i(0) = [0.5x_i, 1.5x_i]$, where x_i are true initial conditions.

Each of the four algorithms was evaluated on each of the 50 datasets and for each ε, we computed ARE and standard deviation as described previously and illustrated in Fig 3. These results show that the Picard method outperforms all methods for low level of noise $\varepsilon \leq 0.1$ and for larger values of noise ($\varepsilon \geq 0.2$) Picard preforms significantly better than SINDy and fmincon methods – but comparable to the proposed black box optimization alternative to the gradient-based Picard approach. Simulation of the Lorenz system and comparison to data for estimated parameters in each method (with noise variance $\varepsilon = 0.4$ and single dataset) is given in Fig. 4.

8 Conclusion

In this paper, we reformulate the problem of parameter estimation for nonlinear ordinary differential equations as a constrained optimization problem with infinite-dimensional variables and constraints. We then propose a new class of gradient-contractive algorithms, based on the contractive properties of the Picard iteration, which eliminates the need for infinite-dimensional constraints and variables (Algs. 2-4). In contrast with existing methods, this approach is gradient-based and allows for data which is sparse, irregular, and includes only partial measurements of the state and does not require numerical simulation or measurements of time-derivatives. Furthermore, we have proposed sufficient conditions under which the proposed algorithm converges to a local optima. Finally, we have exhaustively tested several aspects of performance of the algorithm on a battery of nonlinear models and noisy datasets.

Acknowledgments

We would like to thank Rolf Findeison for his feedback and acknowledge the priority of his team in recognizing the role of Picard iteration in parameter estimation.

References

[1] L. Ljung, "System Identification," in *Signal Analysis and Prediction*. Springer, 1998, pp. 163–173.

[2] N. Oliver, "Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models," 2013.

[3] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2022.

[4] B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," *Proceedings of the National Academy of Sciences*, vol. 17, no. 5, pp. 315–318, 1931.

[5] Y. Han, W. Hao, and U. Vaidya, "Deep Learning of Koopman Representation for Control," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 1890–1895.

[6] E. Yeung, S. Kundu, and N. Hodas, "Learning deep neural network representations for koopman operators of nonlinear dynamical systems," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 4832–4839.

[7] I. Mezic, "Spectral properties of dynamical systems, model reduction and decompositions," *Nonlinear Dynamics*, vol. 41, pp. 309–325, 2005.

[8] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A datadriven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, pp. 1307–1346, 2015.

[9] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

[10] K. Kaheman, J. N. Kutz, and S. L. Brunton, "SINDy-PI: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics," *Proceedings of the Royal Society A*, vol. 476, no. 2242, p. 20200279, 2020.

[11] G. T. Naozuka, H. L. Rocha, R. S. Silva, and R. C. Almeida, "SINDy-SA framework: enhancing nonlinear system identification with sensitivity analysis," *Nonlinear Dynamics*, vol. 110, no. 3, pp. 2589–2609, 2022.

[12] R. Chartrand, "Numerical differentiation of noisy, nonsmooth data," *International Scholarly Research Notices*, vol. 2011, 2011.

[13] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[14] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational physics*, vol. 378, pp. 686–707, 2019.

[15] T. P. Bohlin, *Practical grey-box process identification: theory and applications*. Springer Science & Business Media, 2006.

[16] M. Peifer and J. Timmer, "Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting," *IET Systems Biology*, vol. 1, no. 2, pp. 78–88, 2007.

[17] R. Aster, C. Thurber, and B. Borchers, *Parameter Estimation and Inverse Problems*. Elsevier Science, 2005.

[18] Y. Bard, *Nonlinear Parameter Estimation*. Academic press New York, 1974, vol. 1209.

[19] L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, vol. 34, no. 1, pp. 1–12, 2010.

[20] L. Biegler, J. Damiano, and G. Blau, "Nonlinear parameter estimation: a case study comparison," *AIChE Journal*, vol. 32, no. 1, pp. 29–45, 1986.

[21] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "Approximate solution of the trust region problem by minimization over two-dimensional subspaces," *Mathematical Programming*, vol. 40, no. 1-3, pp. 247–263, 1988.

[22] J. J. Moré, "The Levenberg-Marquardt algorithm: implementation and theory," in *Proceedings of the Biennial Conference held at Dundee, 1977*. Springer, 2006, pp. 105–116.

[23] J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao, "Parameter estimation for differential equations: a generalized smoothing approach," *Journal of the Royal Statistical Society Series B*, vol. 69, no. 5, pp. 741–796, 2007.

[24] G. Hooker, J. O. Ramsay, and L. Xiao, "Collocinfer: collocation inference in differential equation models," *Journal of Statistical Software*, vol. 75, pp. 1–52, 2016.

[25] N. J.-B. Brunel, "Parameter estimation of ODEs via nonparametric estimators," *Electronic Journal of Statistics*, vol. 2, pp. 1242–1267, 2007.

[26] J. Chen and H. Wu, "Efficient local estimation for time-varying coefficients in deterministic dynamic models with applications to HIV-1 dynamics," *Journal of the American Statistical Association*, vol. 103, no. 481, pp. 369–384, 2008.

[27] J. Cao, J. Z. Huang, and H. Wu, "Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations," *Journal of Computational and Graphical Statistics*, vol. 21, no. 1, pp. 42–56, 2012.

[28] V. Duong and A. R. Stubberud, "System identification by genetic algorithm," in *Proceedings of IEEE Aerospace Conference*, vol. 5. IEEE, 2002, pp. 5–5.

[29] S. Sivanandam and S. Deepa, *Genetic Algorithms*. Springer, 2008.

[30] D. Bertsimas and J. Tsitsiklis, "Simulated annealing," *Statistical Science*, vol. 8, no. 1, pp. 10–15, 1993.

[31] B. T. Polyak, *Introduction to Optimization*. New York, Optimization Software,, 1987.

[32] H. Kunze and E. Vrscay, "Solving inverse problems for ordinary differential equations using the Picard contraction mapping," *Inverse Problems*, vol. 15, no. 3, p. 745, 1999.

[33] F. V. Haeseler, N. Rudolph, R. Findeisen, and H. J. Huber, "Parameter estimation for signal transduction networks from experimental time series using picard iteration," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 191–196, 2018.

[34] S. Slavov and T. Tsvetkov, "Picard-Lindelof iterations and multiple shooting method for parameter estimation," *International Journal of Applied Mathematics*, vol. 33, no. 5, p. 919, 2020.

[35] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*. Springer Science & Business Media, 2013.

[36] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley & Sons, 1997.

[37] W. Rudin *et al.*, *Principles of Mathematical Analysis*. McGraw-hill New York, 1964, vol. 3.

[38] L. Ljung and R. Singh, "Version 8 of the Matlab System Identification Toolbox," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 1826–1831, 2012.

[39] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical journal*, vol. 1, no. 6, pp. 445–466, 1961.

[40] K. Grunert, H. Holden, E. R. Jakobsen, and N. C. Stenseth, "Evolutionarily stable strategies in stable and periodically fluctuating populations: The Rosenzweig–Macarthur predator–prey model," *Proceedings of the National Academy of Sciences*, vol. 118, no. 4, 2021.

[41] J. A. Baaijens and J. Draisma, "On the existence of identifiable reparametrizations for linear compartment models," *SIAM Journal on Applied Mathematics*, vol. 76, no. 4, pp. 1577–1605, 2016.

[42] L. G. De Pillis and A. Radunskaya, "The dynamics of an optimally controlled tumor model: A case study," *Mathematical and Computer Modelling*, vol. 37, no. 11, pp. 1221–1244, 2003.

[43] A. Mauroy and J. Goncalves, "Linear identification of nonlinear systems: a lifting technique based on the Koopman operator," in *2016 55th IEEE Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 6500–6505.

[44] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.

A Appendix

In this appendix, we show that gradients of Picard Iterations are Lipschitz continuous. Specifically, we first show that $\nabla_{x,\theta}\mathcal{P}_{t_0,x,\theta}\mathbf{u}$ is Lipschitz continuous with respect to $\mathbf{u}$ uniformly in $(x,\theta) \in X \times \Theta$.

Lemma 17 *Suppose the conditions of Lem. 13 are satisfied. Then, there exists $K_s \geq 0$ such that for any $n \geq m$, $(x,\theta) \in X \times \Theta$ and $\mathbf{u}, \mathbf{v} \in C_a$ we have*

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x,\theta}\mathbf{u} - \nabla_{x,\theta}\mathcal{P}^m_{t_0,x,\theta}\mathbf{v}\|_\infty \leq (TK_x)^m \left[mK_s\|\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u} - \mathbf{v}\|_\infty + \|\nabla_{x,\theta}\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u}\|_\infty \right].$$

PROOF. Define composite functions $\mathbf{u}_n(t,x,\theta) := (\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)$ and $\mathbf{v}_m(t,x,\theta) = (\mathcal{P}^m_{t_0,x,\theta}\mathbf{v})(t)$. Then, for all $x \in X$, $\theta \in \Theta$ and $t \in [0,T)$

$$\|\nabla_x \mathbf{u}_n(t,x,\theta) - \nabla_x \mathbf{v}_m(t,x,\theta)\|_2 = \left\| \int_0^t \nabla_x f(t_0+s,\mathbf{u}_{n-1},\theta)\nabla_x\mathbf{u}_{n-1}(s,x,\theta) - \nabla_x f(t_0+s,\mathbf{v}_{m-1},\theta)\nabla_x\mathbf{v}_{m-1}(s,x,\theta)ds \right\|_2.$$

Note, the Lipschitz continuity of f, $\mathbf{u}_n$, $\mathbf{v}_n$ (from Lem. 11) implies f, $\nabla_{x,\theta}\mathbf{u}_n$, $\nabla_{x,\theta}\mathbf{v}_n$ are bounded.

Next, for some $K_1 \geq 0$ we have for all $x \in X$, $\theta \in \Theta$ and $t \in [0,T)$

$$\|\nabla_x \mathbf{u}_n(t,x,\theta) - \nabla_x \mathbf{v}_m(t,x,\theta)\|_2 \leq K_x \left\| \int_0^t \nabla_x\mathbf{u}_{n-1}(s,x,\theta) - \nabla_x\mathbf{v}_{m-1}(s,x,\theta)ds \right\|_2 + K_1 \left\| \int_0^t \nabla_x f(t_0+s,\mathbf{u}_{n-1},\theta) - \nabla_x f(t_0+s,\mathbf{v}_{m-1},\theta)ds \right\|_2.$$

Thus, for all $x \in X$, $\theta \in \Theta$ and $t \in [0,T)$

$$\sup_{t\in[0,T)} \|\nabla_x \mathbf{u}_n(t,x,\theta) - \nabla_x \mathbf{v}_m(t,x,\theta)\|_2 \leq TK_x \Big(\sup_{t\in[0,T)} \|\nabla_x \mathbf{u}_{n-1}(t,x,\theta) - \nabla_x \mathbf{v}_{m-1}(t,x,\theta)\|_2 + K_1/K_x \sup_{t\in[0,T)} \|\mathbf{u}_{n-1}(t,x,\theta) - \mathbf{v}_{m-1}(t,x,\theta)\|_2 \Big).$$

Therefore, by induction we have for all $x \in X$ and $\theta \in \Theta$ we have

$$\|\nabla_x\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}) - \nabla_x\mathcal{P}^m_{t_0,x,\theta}\mathbf{v}\|_\infty \leq (TK_x)^m \left[mK_s\|\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u} - \mathbf{v}\|_\infty + \|\nabla_x\mathcal{P}^{n-m}_{t_0,x,\theta}\mathbf{u}\|_\infty \right].$$

$\nabla_\theta\mathcal{P}^n_{t_0,x,\theta}\mathbf{u}$ can be considered analogously. ■

Next, we show that $\nabla_{x,\theta}\mathcal{P}_{t_0,x,\theta}\mathbf{u}$ is Lipschitz continuous with respect to x and θ uniformly in $\mathbf{u} \in C_a$.

Lemma 18 *Suppose the conditions of Lem. 13 are satisfied. Then, there exists $K > 0$ such that for any $n \in \mathbb{N}$, $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$ we have*

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x_1,\theta_1}\mathbf{u} - \nabla_{x,\theta}\mathcal{P}^m_{t_0,x_2,\theta_2}\mathbf{u}\|_\infty \leq K\|x_1 - x_2\|_2 + K\|\theta_1 - \theta_2\|_2.$$

PROOF. Define function $\mathbf{u}_n(t,x,\theta) := (\mathcal{P}^n_{t_0,x,\theta}\mathbf{u})(t)$, then for all $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$

$$\|\nabla_x \mathbf{u}_n(t,x_1,\theta_1) - \nabla_x \mathbf{u}_n(t,x_2,\theta_2)\|_2 = \left\| \int_0^t \nabla_x f(t_0+s,\mathbf{u}_{n-1},\theta_1)\nabla_x\mathbf{u}_{n-1}(s,x_1,\theta_1) - \nabla_x f(t_0+s,\mathbf{u}_{n-1},\theta_2)\nabla_x\mathbf{u}_{n-1}(s,x_2,\theta_2)ds \right\|_2$$

Next, since $\nabla_x f$, $\nabla_x\mathbf{u}_{n-1}(s,x_2,\theta_2)$ are bounded (since f and $\mathbf{u}_{n-1}$ are Lipschitz from Lem. 11) there exists $K_1 \geq 0$ such that

$$\|\nabla_x \mathbf{u}_n(t,x_1,\theta_1) - \nabla_x \mathbf{u}_n(t,x_2,\theta_2)\|_2 \leq K_x \left\| \int_0^t \nabla_x\mathbf{u}_{n-1}(s,x_1,\theta_1) - \nabla_x\mathbf{u}_{n-1}(s,x_2,\theta_2)ds \right\|_2 + K_1 \left\| \int_0^t \nabla_x f(t_0+s,\mathbf{u}_{n-1}(t,x_1,\theta_1),\theta_1) - \nabla_x f(t_0+s,\mathbf{u}_{n-1}(t,x_2,\theta_2),\theta_2)ds \right\|_2$$

for all $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$.

Thus, since $\nabla_x f$ and $\mathbf{u}_{n-1}$ is Lipschitz continuous in x and θ there exists $K_2 \geq 0$ such that

$$\sup_{t\in[0,T)} \|\nabla_x \mathbf{u}_n(t,x_1,\theta_1) - \nabla_x \mathbf{u}_n(t,x_2,\theta_2)\|_2 \leq TK_x \sup_{t\in[0,T)} \|\nabla_x \mathbf{u}_{n-1}(t,x_1,\theta_1) - \nabla_x \mathbf{u}_{n-1}(t,x_2,\theta_2)\|_2 + K_2\|x_1 - x_2\| + K_2\|\theta_2 - \theta_1\|_2.$$

for all $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$.

Next, since $\mathbf{u} \in C_a$ is not a function of x and θ, by induction for all $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$ we have

$$\sup_{t\in[0,T)} \|\nabla_x \mathbf{u}_n(t,x_1,\theta_1) - \nabla_x \mathbf{u}_n(t,x_2,\theta_2)\|_2 \leq \frac{K_2}{1-TK_x}\|x_1 - x_2\| + \frac{K_2}{1-TK_x}\|\theta_2 - \theta_1\|_2.$$

Analogously, we have that $\nabla_p\mathbf{u}_n(t,x,p)$ is Lipschitz continuous. Thus, for $K = \frac{K_2}{1-TK_x}$ we have

$$\|\nabla_{x,\theta}\mathcal{P}^n_{t_0,x_1,\theta_1}\mathbf{u} - \nabla_{x,\theta}\mathcal{P}^m_{t_0,x_2,\theta_2}\mathbf{u}\|_\infty \leq K\|x_1 - x_2\|_2 + K\|\theta_1 - \theta_2\|_2$$

for any $n \in \mathbb{N}$, $(x_1,\theta_1),(x_2,\theta_2) \in X \times \Theta$ and $\mathbf{u} \in C_a$. ■