A LIMITED-MEMORY MULTIPLE SHOOTING METHOD FOR WEAKLY CONSTRAINED VARIATIONAL DATA ASSIMILATION[§]

WANTING XU[†] AND MIHAI ANITESCU[‡]

Abstract. We present a limited-memory method for maximum-likelihood-based state estimation of hidden Markov models. We reduce the memory storage requirements by expressing the optimal states as a function of checkpoints bounding a shooting interval. All states can then be recomputed as needed from a recursion stemming from the optimality conditions. The matching of states at checkpoints are imposed, in a multiple shooting fashion, as constraints on the optimization problem, which is solved with an augmented Lagrangian method. We prove that for nonlinear systems under certain assumptions the condition number of the Hessian matrix of the augmented Lagrangian function is bounded above with respect to the number of shooting intervals. Hence the method is stable for increasing time horizon. The assumptions include satisfying the observability conditions of the linearized system on a shooting interval. We also propose a recursion-based gradient evaluation algorithm for computing the gradient, which in turn allows the algorithm to proceed by storing at any time only the checkpoints and the states on a shooting interval. We demonstrate our findings with simulations in different regimes for Burgers' equation.

Key words. data assimilation, hidden Markov models, multiple shooting, observability, stability

AMS subject classifications. 62M99, 93B99

1. Introduction. Data assimilation is the process of estimating the underlying states of a physical system based on reconciliation of observations and physical laws governing its evolution [4, 6, 13]. The setup is most commonly described by a hidden Markov model with stochastic normal model error and measurement noise [13],

$$x_0 = x_B + \eta_B, \;\; x_{j+1} = M_j(x_j) + \eta_j, \;\; y_j = H_j(x_j) + \varepsilon_j, \tag{1.1}$$

$$\eta_B \sim \mathcal{N}(\mathbf{0}_J, Q_B), \;\; \eta_j \sim \mathcal{N}(\mathbf{0}_J, Q_j) \;\; \varepsilon_j \sim \mathcal{N}(\mathbf{0}_L, R_j). \tag{1.2}$$

where $x_j \in \mathbb{R}^J, y_j \in \mathbb{R}^L$. The mapping $M_j(\cdot) : \mathbb{R}^J \to \mathbb{R}^J$ models the physical law governing the evolution of the system dynamics, typically discretizations of partial differential equations. We assume $M_j(\cdot)$ is at least twice continuously differentiable. The random variable η_j models the stochastic model error and has a normal distribution with mean $\mathbf{0}_J$ and covariance $Q_j \in \mathbb{R}^{J\times J}$. The random variable η_B models the initial state as a normal distribution with mean x_B and covariance $Q_B \in \mathbb{R}^{J\times J}$. The function $H_j(\cdot) : \mathbb{R}^J \to \mathbb{R}^L$ maps the states into observed quantities, whereas ε_j models measurement error that has mean $\mathbf{0}_L$ and covariance $R_j \in \mathbb{R}^{L\times L}$. We also assume all covariance matrices to be positive definite.

With these definitions, we are interested in the *state estimation problem* [23]: We are given the background mean state x_B; evolution function $M_j(\cdot)$; measurement operator $H_j(\cdot)$; measured quantities y_j; and covariance matrices for background error, Q_B, model error, Q_j for $j = 0, 1, \ldots, N-1$, and measurement error, R_j for $j = 0, 1, \ldots, N$ at $N+1$ equally spaced time points. We want to determine the state trajectory $x_0, x_1, \ldots, x_N$ that best explains the data y_j under these assumptions. The problem is also named *data assimilation* or *4DVar* [4, 6, 13, 17] in atmospheric sciences applications, when $M_j(\cdot)$ is obtained from the discretization of 3D dynamics.

[†]Department of Statistics, University of Chicago, Chicago, IL 60637 (wxu@galton.uchicago.edu).

[‡]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439 (anitescu@mcs.anl.gov).

[§]Also, Preprint, ANL/MCS-P5446-1215

In particular, we will focus on the circumstance where we are memory-limited, and thus we may be unwilling to simultaneously store the entire trajectory vector because of the $O(JN)$ memory requirements.

In the limiting case of $Q_j = \mathbf{0}_{J\times J}$, and thus $\eta_j = \mathbf{0}_J$, the system is called "strongly constrained" in the sense that every state is determined by the previous one and all states are functions of only the initial state x_0. However, many sources (e.g., discretization and parametrization) can contribute to model errors that have non-negligible effects [5, 24, 25]. The explicit inclusion of the model error term in the physical evolution [8, 19, 20] leads precisely to (1.1)–(1.2). This paradigm is called a *hidden Markov model* [21, 22], and it is one of the most studied state estimation paradigms [13]. It has generated a large number of methods to solve it, including Kalman filters, extended Kalman filters, and particle methods [7, 13]. However, such methods may not be suitable to the kind of problems described here because of reliance on linearity of $M_j(\cdot)$ (Kalman filters) [12]; memory that increases superlinearly with the dimension of x (extended Kalman filters) [10]; and slow convergence, particularly when interested primarily in best estimates (particle methods) [3].

In this work we focus on *variational methods*: methods that aim to express the minus loglikelihood of the hidden Markov model (1.1)–(1.2) and then minimize it with deterministic methods, such as limited-memory BFGS [18]. The objective function of that minimization is the following weakly constrained function [15, 16, 25, 26, 27]:

$$\Gamma(x_{0:N}) = \frac{1}{N}\left(\sum_{j=0}^{N-1}\left(\gamma_j(x_j) + \phi_j\left(x_j, x_{j+1}\right)\right) + \gamma_N(x_N)\right), \tag{1.3}$$

where
$$\phi_j(x_j, x_{j+1}) = \left(x_{j+1} - M_j(x_j)\right)^T Q_j^{-1}\left(x_{j+1} - M_j(x_j)\right)/2, \qquad 0 \le j \le N-1$$
$$\gamma_j(x_j) = \left(y_j - H_j(x_j)\right)^T R_j^{-1}\left(y_j - H_j(x_j)\right)/2, \qquad 1 \le j \le N$$
$$\gamma_0(x_0) = (x_0 - x_B)^T Q_B^{-1}(x_0 - x_B)/2 + \left(y_0 - H_0(x_0)\right)^T R_0^{-1}\left(y_0 - H_0(x_0)\right)/2.$$

The best estimation of the states $x_0, x_1, \ldots, x_N$ then amounts to minimizing (1.3), which is equivalent to maximizing the likelihood of the hidden Markov model. In the strongly constrained case, only x_0 is a free variable. Using adjoint approaches for the minimization of (1.3) in that limiting case with a checkpointing strategy results in storage requirements of about $O(J\log(N))$ with a recomputation effort that is relatively bounded with J and N [9]. In the presence of model error, however, it is no longer possible to constrain the states by using model propagation, and hence the storage is $N+1$ fold larger since all states $x_0, x_1, \ldots, x_N$ are free variables. In the case of a large J or N, which we are increasingly approaching in atmospheric sciences as more refined physics models are coming online, the sheer amount of storage makes applications to real systems with higher resolution out of practical reach.

To this end, we recently [1] proposed to reduce memory by using the constraints of the optimality conditions themselves.

$$0 = \nabla_{x_0}\phi_0(x_0, x_1) + \nabla_{x_0}\gamma_0(x_0) \tag{1.4}$$
$$0 = \nabla_{x_j}\phi_j(x_j, x_{j+1}) + \nabla_{x_j}\phi_{j-1}(x_{j-1}, x_j) + \nabla_{x_j}\gamma_j(x_j), \qquad 1 \le j \le N \tag{1.5}$$
$$0 = \nabla_{x_N}\phi_{N-1}(x_{N-1}, x_N) + \nabla_{x_N}\gamma_N(x_N) \tag{1.6}$$

Enforcing optimality conditions (1.4) and (1.5) gives a recursion for computing x_1 in terms of x_0 and x_{i+1} in terms of x_i and x_{i-1} for $1 \le i \le N-1$. Hence each state effectively is reduced to a function of just the initial state by using the optimality

conditions as constraints; we call this recursively computable function $\lambda_i(x_0)$, $i = 1, 2, \ldots, N$. The objective function then becomes

$$\hat{\Gamma}(x_0) = \frac{1}{N}\left(\sum_{i=0}^{N-1} \gamma_i\left(\lambda_i(x_0)\right) + \phi_i\left(\lambda_i(x_0), \lambda_{i+1}(x_0)\right) + \gamma_N\left(\lambda_N(x_0)\right)\right). \tag{1.7}$$

Quasi-Newton methods such as L-BFGS can be used to minimize (1.7).

The recursive nature of the method opens the door for instability when the time horizon increases or under certain model parameters, as also discussed in [1]. That is, the recursion may exhibit rapid exponential increase of the solution, resulting in numerical overflow. Numerical experiments show that in the presence of large model error, large observation gap, large time step, or increased time horizon, the method may encounter such stability issues and fail to progress. The method that minimizes (1.7) in [1] uses essentially a single shooting idea. Each initial state x_0 determines the whole trajectory through $\lambda_i(x_0)$, and the optimality is found by satisfying the optimality condition at the end $\nabla_{x_N}\phi_{N-1} + \nabla_{x_N}\gamma_N = 0$. We propose a multiple shooting approach for which multiple restart points across the whole horizon are used. We call such restart points *checkpoints*, given their identical functionality in adjoint calculations [9]. Each checkpoint sequence determines a "shooting" segment of the trajectory, and optimality is achieved by both minimizing the resulting function and matching at each checkpoint. To compute the function and its gradients on a shooting interval, we use a recursion like (1.5) restarted at the last checkpoint; a "shooting" approach. At the cost of modestly increased storage, we expect the method to improve stability by reducing the length of recursion on each segment.

The rest of this article is organized as follows. Section 2 describes the low-memory multiple shooting method and proves the consistency of the solution with the full-memory data assimilation method. In Section 3, we show that for nonlinear systems within a certain regime, the condition number of the multiple shooting method is bounded above with respect to the number of shooting intervals. Section 4 describes a recursive limited-memory algorithm to evaluate the descent direction of the resulting optimization problem in preparation for numerical experiments. Section 5 presents numerical experiments that implement the multiple shooting method for Burgers' equation under different parameter settings. Improvements and limitations are discussed in the conclusion.

2. Multiple shooting approach. We note that the recursion defining x_{j+1} through (1.5) is a two-term recursion; therefore a checkpointing approach here would need two consecutive states. In the following, d pairs of checkpoints $\{x_{P_1-1}, x_{P_1}, \ldots, x_{P_d-1}, x_{P_d}\} \in \mathbb{R}^{2dJ}$ are equally spaced across the entire state. To simplify the discussion, we assume that the number of states on each shooting interval is constant; we let $k = N/(d+1)$ be that number. We also denote $P_0 = 0$ and $P_{d+1} = N$. For each shooting interval $[x_{P_i}, x_{P_{i+1}}]$ we define by $\hat{\Gamma}_i$ the component of the objective function (1.3) attached to that interval:

$$\hat{\Gamma}_0(x_0) = \frac{1}{N}\left(\sum_{j=0}^{P_1-1} \gamma_j(\widetilde{x}_j(x_0)) + \phi_j(\widetilde{x}_j(x_0), \widetilde{x}_{j+1}(x_0))\right), \tag{2.1a}$$

$$\hat{\Gamma}_i(x_{P_i-1}, x_{P_i}) = \frac{1}{N}\left(\sum_{j=P_i}^{P_{i+1}-1} \gamma_j(\widetilde{x}_j(x_{P_i-1}, x_{P_i}))\right. \tag{2.1b}$$

$$+\phi_j(\widetilde{x}_j(x_{P_i-1},x_{P_i}),\widetilde{x}_{j+1}(x_{P_i-1},x_{P_i}))\Bigg), \qquad 1\le i\le d-1$$

$$\hat{\Gamma}_d(x_{P_d-1},x_{P_d}) = \frac{1}{N}\Bigg(\sum_{j=P_d}^{N-1}\gamma_k(\widetilde{x}_j(x_{P_i-1},x_{P_i})) \tag{2.1c}$$

$$+\phi_j(\widetilde{x}_j(x_{P_i-1},x_{P_i}),\widetilde{x}_{j+1}(x_{P_i-1},x_{P_i}))+\gamma_N(\widetilde{x}_N(x_{P_d}))\Bigg).$$

The mappings $\widetilde{x}_j(x_{P_i-1},x_{P_i})$ are defined implicitly from the optimality conditions (1.4) and (1.5). This step is possible as soon as $\nabla_{x_j}\phi(x_j,x_{j+1}) = \nabla_{x_j}M_j(x_j)Q_j^{-1}(x_{j+1}-M_j(x_j))$ is invertible in x_{j+1}. This is equivalent to requiring that $\nabla_{x_j}M_j(x_j)Q_j^{-1}$ be an invertible matrix. Since $M_j(\cdot)$ are propagating operators, they can be assumed to be invertible from properties of dynamical systems (see also the discussion at the beginning of [1, §3]). Since the covariance matrix Q_j is assumed to be positive definite, it immediately follows that the recursion (1.5) is uniquely solvable in x_{j+1}.

At points immediately following the checkpoints, the mappings $\widetilde{x}_{P_i+1}(x_{P_i-1},x_{P_i})$ are the solution of the optimality conditions (1.4) and (1.5) at checkpoint P_i:

$$0 = \nabla_{x_0}\gamma_{x_0}(x_0)+\nabla_{x_0}\phi_{P_0}(x_0,\widetilde{x}_1) \tag{2.2}$$

$$0 = \nabla_{x_{P_i}}\phi_{P_{i-1}}(x_{P_i-1},x_{P_i})+\nabla_{x_{P_i}}\gamma_{P_i}(x_{P_i})+\nabla_{x_{P_i}}\phi_{P_i}(x_{P_i},\widetilde{x}_{P_i+1}), \tag{2.3}$$

for $i=1,\dots,d$. At all other points, $\widetilde{x}_j(x_{P_i-1},x_{P_i})$ is defined recursively from $\widetilde{x}_{j-1}(x_{P_i-1},x_{P_i})$ and $\widetilde{x}_{j-2}(x_{P_i-1},x_{P_i})$ by using the optimality conditions (1.5) as follows:

$$0 = \nabla_{x_j}\phi_{j-1}(\widetilde{x}_{j-1},\widetilde{x}_j)+\nabla_{x_j}\gamma_j(\widetilde{x}_j)+\nabla_{x_j}\phi_j(\widetilde{x}_j,\widetilde{x}_{j+1}), \tag{2.4}$$

for $P_i < j \le P_{i+1}-1$, $i=0,\dots,d$. Under model (1.1), the recursions (2.2)–(2.4) can be written at points immediately following checkpoints as

$$\begin{aligned}\widetilde{x}_1(x_0) &= M_0(x_0)+Q_0\nabla^{-T}M_0(x_0)Q_B^{-1}(x_0-x_B)\\ &\quad -Q_0\nabla^{-T}M_0(x_0)\nabla^T H_0(x_0)R_0^{-1}(y_0-H_0(x_0)),\end{aligned} \tag{2.5}$$

$$\begin{aligned}\widetilde{x}_{P_i+1}(x_{P_i},x_{P_i-1}) &= M_{P_i}(x_{P_i})+Q_{P_i}\nabla^{-T}M_{P_i}(x_{P_i})Q_{P_i-1}^{-1}(x_{P_i}-M_{P_i-1}(x_{P_i-1}))\\ &\quad -Q_{P_i}\nabla^{-T}M_{P_i}(x_{P_i})\nabla^T H_{P_i}(x_{P_i})R_{P_i}^{-1}(y_{P_i}-H_{P_i}(x_{P_i})),\end{aligned} \tag{2.6}$$

for $i=1,2,\dots,d$. At all other points between checkpoints we obtain

$$\begin{aligned}\widetilde{x}_{j+1}(\widetilde{x}_j,\widetilde{x}_{j-1}) &= M_j(\widetilde{x}_j)+Q_j\nabla^{-T}M_j(\widetilde{x}_j)Q_{j-1}^{-1}(\widetilde{x}_j-M_{j-1}(\widetilde{x}_{j-1}))\\ &\quad -Q_j\nabla^{-T}M_j(\widetilde{x}_j)\nabla^T H_j(\widetilde{x}_j)R_j^{-1}(y_j-H_j(\widetilde{x}_j)).\end{aligned} \tag{2.7}$$

Repeated use of (2.7) together with (2.5) and (2.6) results in computing all mappings $\widetilde{x}_j(x_{P_i-1},x_{P_i})$

Then, by gathering the objective function components (2.1) and by imposing matching constraints at the checkpoint pairs, we obtain the following multiple shooting optimization problem.

$$\min \quad \widetilde{\Gamma}(x_0,x_{P_1-1},x_{P_1},\dots,x_{P_d-1},x_{P_d}) \triangleq \hat{\Gamma}_0(x_0)+\sum_{i=1}^{d}\hat{\Gamma}_i(x_{P_i-1},x_{P_i}) \tag{2.8a}$$

$$\text{(2.8b)} \quad \text{s.t.} \quad c_1(x) = x_{P_1} - \widetilde{x}_{P_1}(x_0) = 0$$
$$\text{(2.8c)} \quad g_1(x) = x_{P_1-1} - \widetilde{x}_{P_1-1}(x_0) = 0$$
$$\text{(2.8d)} \quad c_{i+1}(x) = x_{P_{i+1}} - \widetilde{x}_{P_{i+1}}(x_{P_i-1}, x_{P_i}) = 0, \qquad 1 \le i \le d-1$$
$$\text{(2.8e)} \quad g_{i+1}(x) = x_{P_{i+1}-1} - \widetilde{x}_{P_{i+1}-1}(x_{P_i-1}, x_{P_i}) = 0, \qquad 1 \le i \le d-1$$

The Lagrangian associated with the constraint problem (2.8) is

$$L(x, \lambda, \psi) = \widetilde{\Gamma}(x) - \sum_{i=1}^{d} \lambda_i^T c_i(x) - \sum_{i=1}^{d} \psi_i^T g_i(x), \tag{2.9}$$

where $x = (x_0, x_{P_1-1}, x_{P_1}, \ldots, x_{P_d-1}, x_{P_d})$ and $\lambda_i \in \mathbb{R}^J$, $\psi_i \in \mathbb{R}^J$ are Lagrange multipliers for the equality constraints $c_i(x) = 0$ and $g_i(x) = 0$, $i = 1, 2, \ldots, d$.

We also define the full memory form of the objective functions for each shooting interval as follows:

$$\begin{aligned} \Gamma_i(x_{P_i:P_{i+1}}) &= \frac{1}{N} \left(\sum_{j=P_i}^{P_{i+1}-1} \gamma_j(x_k) + \phi_j(x_j, x_{j+1}) \right), \qquad 0 \le i \le d, \\ \Gamma_d(x_{P_d:N}) &= \frac{1}{N} \left(\sum_{j=P_d}^{N-1} \gamma_j(x_j) + \phi_j(x_j, x_{j+1}) + \gamma_N(x_N) \right). \end{aligned} \tag{2.10}$$

We now define a list of symbols frequently used in the rest of the article.

DEFINITION 2.1. *For $1 \le i \le d$ and $0 \le j \le N$, define*

(a)

$$\begin{aligned} \beta_j(x_j, x_{j+1}) &= \nabla_{x_j} \gamma_j(x_j) + \nabla_{x_j} \phi_j(x_j, x_{j+1}), \qquad 0 \le j \le N-1 \\ \alpha_j(x_{j-1}, x_j) &= \nabla_{x_j} \phi_{j-1}(x_{j-1}, x_j), \qquad 1 \le j \le N \\ \theta_j(x_{j-1}, x_j, x_{j+1}) &= \alpha_j(x_{j-1}, x_j) + \beta_j(x_j, x_{j+1}), \qquad 1 \le j \le N-1 \\ \theta_0(x_0, x_1) &= \beta_0(x_0, x_1); \;\; \theta_N(x_{N-1}, x_N) = \alpha_N(x_{N-1}, x_N) + \nabla_{x_N} \gamma_N(x_N) \end{aligned}$$

Note that for Γ_i defined in (2.10), we have

$$\left(\frac{\partial \Gamma_i}{\partial (x_{P_i:P_{i+1}})} \right)^T = \left[\beta_{P_i}^T, \theta_{P_i+1}^T, \ldots, \theta_{P_{i+1}-1}^T, \alpha_{P_{i+1}}^T \right], \qquad 0 \le i \le d-1$$
$$\left(\frac{\partial \Gamma_d}{\partial (x_{P_d:N})} \right)^T = \left[\beta_{P_d}^T, \theta_{P_d+1}^T, \ldots, \theta_{N-1}^T, \theta_N^T \right].$$

(b)

$$\begin{aligned} L_j^{(0)}(x_0) &= \nabla_{x_0} \widetilde{x}_j(x_0), \qquad 0 \le j \\ L_j^{(P_i-1)}(x_{P_i-1}, x_{P_i}) &= \nabla_{x_{P_i-1}} \widetilde{x}_j(x_{P_i-1}, x_{P_i}), \qquad P_i - 1 \le j \\ L_j^{(P_i)}(x_{P_i-1}, x_{P_i}) &= \nabla_{x_{P_i}} \widetilde{x}_j(x_{P_i-1}, x_{P_i}), \qquad P_i - 1 \le j \end{aligned}$$

(c) Let $\Lambda_i(x_{P_i-1}, x_{P_i})$ be $(k+1)J \times 2J$ dimensional, and let $\Lambda_0(x_0)$ be $(k+1)J \times J$ dimensional matrices so that

$$\Lambda_i(x_{P_i-1}, x_{P_i}) = \frac{\partial (\widetilde{x}_{P_i:P_{i+1}})}{\partial (x_{P_i-1}, x_{P_i})} = \begin{bmatrix} L_{P_i}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_i}^{(P_i)}(x_{P_i-1}, x_{P_i}) \\ L_{P_i+1}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_i+1}^{(P_i)}(x_{P_i-1}, x_{P_i}) \\ \vdots & \vdots \\ L_{P_{i+1}}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_{i+1}}^{(P_i)}(x_{P_i-1}, x_{P_i}) \end{bmatrix},$$

$$\Lambda_0(x_0) = \frac{\partial(\widetilde{x}_{0:P_1})}{\partial(x_0)} = \left[L_0^{(0)}(x_0)^T, L_1^{(0)}(x_0)^T, \ldots, L_{P_1}^{(0)}(x_0)^T\right]^T.$$

Note that the first block row of Λ_i is $[0, I_J]$ and the first block row of Λ_0 is I_J. Let $L_0(x_0)$ and $L_i(x_{P_i-1}, x_{P_i})$ be the last two block rows respectively of $\Lambda_0(x_0)$ and $\Lambda_i(x_{P_i-1}, x_{P_i})$ so that

$$L_0(x_0) = \begin{bmatrix} L_{P_1-1}^{(0)}(x_0) \\ L_{P_1}^{(0)}(x_0) \end{bmatrix}$$

$$L_i(x_{P_i-1}, x_{P_i}) = \begin{bmatrix} L_{P_{i+1}-1}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_{i+1}-1}^{(P_i)}(x_{P_i-1}, x_{P_i}) \\ L_{P_{i+1}}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_{i+1}}^{(P_i)}(x_{P_i-1}, x_{P_i}) \end{bmatrix}.$$

(d) Let $J_i(x_{P_i-1}, x_{P_i})$ and $J_0(x_0)$ be $J(k+1) \times J(k+1)$ dimensional symmetric block tridiagonal matrices defined as follows (with the arguments of $\beta_.$,$\theta_.$,$\alpha_.$ dropped for brevity).

$$J_i = \begin{bmatrix} \nabla_{x_{P_i}} \beta_{P_i} & \nabla_{x_{P_i+1}} \theta_{P_i} & & \mathbf{0} \\ \nabla_{x_{P_i}} \theta_{P_i+1} & \nabla_{x_{P_i+1}} \theta_{P_i+1} & \ddots & \\ & \ddots & \ddots & \\ & \ddots & \nabla_{x_{P_{i+1}-1}} \theta_{P_{i+1}-1} & \nabla_{x_{P_{i+1}}} \theta_{P_{i+1}-1} \\ \mathbf{0} & & \nabla_{x_{P_{i+1}-1}} \theta_{P_{i+1}} & \nabla_{x_{P_{i+1}}} \alpha_{P_{i+1}} \end{bmatrix}.$$

Note that $J_i = \nabla^2 \Gamma_i$ for $0 \le i \le d-1$, and $\nabla^2 \Gamma_d$ differs from J_d by only the last diagonal block element so that $(J_d)_{(k,k)} + \nabla^2_{x_N} \gamma_N = (\nabla^2 \Gamma_d)_{(k,k)}$.

We now illustrate the relationship between the solution of the multiple shooting constrained optimization problem (2.8) and the solution of the full-memory data assimilation problem (1.3)

THEOREM 2.2. *Let $x^*_{0:N}$ be a local minimizer of $\Gamma(x_{0:N})$ (1.3) that satisfies the first- and second-order sufficient conditions. Let $x^* = (x^*_0, x^*_{P_1-1}, x^*_{P_1}, \ldots, x^*_{P_d-1}, x^*_{P_d})$. Then*

(a) x^* satisfies the KKT conditions of (2.8) with Lagrangian multipliers $\lambda^*_i = -\nabla_{x_{P_i}} \phi_{P_i-1}(x^*_{P_i-1}, x^*_{P_i})$, $\psi^*_i = 0$ for $1 \le i \le d$.

(b) The Hessian matrix of the Lagrangian at optimality satisfies

$$w^T \nabla^2_x L(x^*, \lambda^*, \psi^*) w = \sum_{i=0}^{d} \hat{w}_i^T \Lambda_i^T J_i \Lambda_i \hat{w}_i + \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right)^T \nabla^2_{x_N} \gamma_N \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right),$$

for $w = (w_1, \ldots, w_{2d+1}) \in \mathbb{R}^{(2d+1)J}$, $\hat{w}_i = (w_{2i}, w_{2i+1})$, $1 \le i \le d$, and $\hat{w}_0 = w_1$.

(c) x^* satisfies the second-order sufficient conditions of (2.8).

Proof. The optimality conditions (2.2)–(2.4) uniquely determine the recursion of $\widetilde{x}_j$, $0 \le j \le N$ (Theorem 1 of [1]). Therefore the solution $x^*_{0:N}$ of (1.3) coincides with the state propagated starting from the checkpoints by using the recursions (2.2)–(2.4), namely, $\widetilde{x}_j = x^*_j$ for $0 \le j \le N$. In the rest of the proof, the dependence of the symbols defined in Definition 2.1 on the checkpoints is suppressed for brevity.

First, we aim to verify part (a), that is, check the KKT conditions with Lagrangian multipliers $\lambda_i^* = -\nabla_{x_{P_i}} \phi_{P_i-1}(x^*_{P_{i-1}}, x^*_{P_i})$, $\psi_i^* = 0$ for $1 \le i \le d$. Note that from the definitions of α_{P_i}, β_{P_i} (Definition 2.1(a)) and optimality conditions (1.4) and (1.5), we have that for $1 \le i \le d$,

$$\alpha_{P_i}(x^*_{P_i-1}, x^*_{P_i}) + \lambda_i^* = 0, \tag{2.11a}$$

$$\beta_{P_i}(x^*_{P_i}, x^*_{P_i+1}) - \lambda_i^* = 0. \tag{2.11b}$$

By the chain rule and from the definition of the constraints (2.8b) and (2.8c) and Definitions 2.1(a) and (c), the first-order derivatives are

$$\begin{aligned} \nabla_{x_0} L(x^*, \lambda^*, \psi^*) &= \nabla_{x_0} \hat{\Gamma}_0(x_0^*) - \nabla_{x_0} c_1(x^*)\lambda_1^* - \nabla_{x_0} g_1(x^*)\psi_1^* \\ &= \left(\frac{\partial(\widetilde{x}_{0:P_1})}{\partial(x_0)}\right)^T \frac{\partial \Gamma_0}{\partial(x_{0:P_1})} + {L_{P_1}^{(0)}}^T \lambda_1^* + {L_{P_1-1}^{(0)}}^T \psi_1^* \\ &= \Lambda_0^T V_0 + {L_{P_1-1}^{(0)}}^T \psi_1^*, \text{ where} \end{aligned} \tag{2.12}$$

$$V_0 := \begin{bmatrix} \theta_0(x_0^*, \widetilde{x}_1) \\ \theta_1(\widetilde{x}_0, \widetilde{x}_1, \widetilde{x}_2) \\ \vdots \\ \theta_{P_1-1}(\widetilde{x}_{P_1-2}, \widetilde{x}_{P_1-1}, \widetilde{x}_{P_1}) \\ \alpha_{P_1}(\widetilde{x}_{P_1-1}, \widetilde{x}_{P_1}) + \lambda_i^* \end{bmatrix}. \tag{2.13}$$

Optimality conditions (1.4), (1.5) and (2.11a) imply $V_0 = \mathbf{0}$, and hence we have $\nabla_{x_0} L(x^*, \lambda^*, \psi^*) = \mathbf{0}$.

For $1 \le i \le d-1$, from the definition of the constraints (2.8d) and (2.8e) and Definitions 2.1(a) and (c), we obtain that

$$\begin{aligned} &\nabla_{(x_{P_i-1}, x_{P_i})} L(x^*, \lambda^*, \psi^*) = \nabla_{(x_{P_i-1}, x_{P_i})} \hat{\Gamma}_i(x^*_{P_i-1}, x^*_{P_i}) \\ &- \begin{bmatrix} \nabla_{x_{P_i-1}} g_i(x^*)\psi_i^* + \nabla_{x_{P_i-1}} c_{i+1}(x^*)\lambda_{i+1}^* + \nabla_{x_{P_i-1}} g_{i+1}(x^*)\psi_{i+1}^* \\ \nabla_{x_{P_i}} c_i(x^*)\lambda_i^* + \nabla_{x_{P_i}} c_{i+1}(x^*)\lambda_{i+1}^* + \nabla_{x_{P_i}} g_{i+1}(x^*)\psi_{i+1}^* \end{bmatrix} \\ &= \left(\frac{\partial(\widetilde{x}_{P_i:P_{i+1}})}{\partial(x_{P_i-1}, x_{P_i})}\right)^T \frac{\partial \Gamma_i}{\partial(x_{P_i:P_{i+1}})} - \begin{bmatrix} \psi_i^* - {L_{P_{i+1}-1}^{(P_i-1)}}^T \psi_{i+1}^* - {L_{P_{i+1}}^{(P_i-1)}}^T \lambda_{i+1}^* \\ \lambda_i^* - {L_{P_{i+1}-1}^{(P_i)}}^T \psi_{i+1}^* - {L_{P_{i+1}}^{(P_i)}}^T \lambda_{i+1}^* \end{bmatrix} \\ &= \Lambda_i^T V_i - \begin{bmatrix} \psi_i^* - {L_{P_{i+1}-1}^{(P_i-1)}}^T \psi_{i+1}^* \\ -{L_{P_{i+1}-1}^{(P_i)}}^T \psi_{i+1}^* \end{bmatrix}, \text{ where} \end{aligned} \tag{2.14}$$

$$V_i := \begin{bmatrix} \beta_{P_i}(x^*_{P_i}, \widetilde{x}_{P_i+1}) - \lambda_i^* \\ \theta_{P_i+1}(\widetilde{x}_{P_i}, \widetilde{x}_{P_i+1}, \widetilde{x}_{P_i+2}) \\ \vdots \\ \theta_{P_{i+1}-1}(\widetilde{x}_{P_{i+1}-2}, \widetilde{x}_{P_{i+1}-1}, \widetilde{x}_{P_{i+1}}) \\ \alpha_{P_{i+1}}(\widetilde{x}_{P_{i+1}-1}, \widetilde{x}_{P_{i+1}}) + \lambda_{i+1}^* \end{bmatrix}. \tag{2.15}$$

Optimality conditions (1.4) and (1.5) and (2.11a) and (2.11b) imply that $V_i = \mathbf{0}$, and hence we have $\nabla_{(x_{P_i-1}, x_{P_i})} L(x^*, \lambda^*, \psi^*) = \mathbf{0}$.

For the last shooting interval, from the definition of the constraints (2.8d) and (2.8e) and Definitions 2.1(a) and (c), we obtain that

$$\nabla_{(x_{P_d-1}, x_{P_d})} L(x^*, \lambda^*, \psi^*) = \nabla_{(x_{P_d-1}, x_{P_d})} \hat{\Gamma}_d(x^*_{P_d-1}, x^*_{P_d}) \tag{2.16}$$

$$- \begin{bmatrix} \nabla_{x_{P_d-1}} g_d(x^*)\psi_d^* \\ \nabla_{x_{P_d}} c_d(x^*)\lambda_d^* \end{bmatrix} = \Lambda_d^T V_d - \begin{bmatrix} \psi_d^* \\ \mathbf{0} \end{bmatrix}, \text{ where}$$

$$V_d := \begin{bmatrix} \beta_{P_d}(x_{P_d}^*, \widetilde{x}_{P_d+1}) - \lambda_d^* \\ \theta_{P_d+1}(\widetilde{x}_{P_d}, \widetilde{x}_{P_d+1}, \widetilde{x}_{P_d+2}) \\ \vdots \\ \theta_{N-1}(\widetilde{x}_{N-2}, \widetilde{x}_{N-1}, \widetilde{x}_N) \\ \theta_N(\widetilde{x}_{N-1}, \widetilde{x}_N) \end{bmatrix}. \tag{2.17}$$

Optimality conditions (1.5) and (1.6) and (2.11b) imply that $V_d = \mathbf{0}$, and hence we have $\nabla_{(x_{P_d-1}, x_{P_d})} L(x^*, \lambda^*, \psi^*) = \mathbf{0}$. This completes the proof of part (a).

We now derive the Hessian matrix. For $1 \le i \le d$, directly applying the chain rule to (2.12) and (2.14), we note that $V_i = \mathbf{0}$ for $0 \le i \le d-1$ give that $\nabla^2_{x_0} L(x^*, \lambda^*, \psi^*) = \Lambda_0^T J_0 \Lambda_0$ and that $\nabla^2_{(x_{P_i-1}, x_{P_i})} L(x^*, \lambda^*, \psi^*) = \Lambda_i^T J_i \Lambda_i$ for $1 \le i \le d-1$.

For the last shooting interval, applying the chain rule to (2.16) and from Definitions 2.1(a) and (d) and the fact that $V_d = \mathbf{0}$, we obtain that

$$\nabla^2_{(x_{P_d-1}, x_{P_d})} L(x^*, \lambda^*, \psi^*) = \Lambda_d^T J_d \Lambda_d + \begin{bmatrix} {L_N^{(P_d-1)}}^T \\ {L_N^{(P_d)}}^T \end{bmatrix} \nabla^2_{x_N} \gamma_N \left[L_N^{(P_d-1)} L_N^{(P_d)} \right].$$

Since the constraints are separable, there are no cross terms in the Hessian matrix.

For $w = (w_1, \ldots, w_{2d+1}) \in \mathbb{R}^{(2d+1)J}$, we define $\hat{w}_i = (w_{2i}, w_{2i+1})$ for $1 \le i \le d$ and $\hat{w}_0 = w_1$. Then we have that

$$\begin{aligned} w^T \nabla_x^2 L(x^*, \lambda^*, \mu) w &= \sum_{i=0}^{d} \hat{w}_i^T \Lambda_i^T J_i \Lambda_i \hat{w}_i \\ &+ \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1} \right)^T \nabla^2_{x_N} \gamma_N \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1} \right). \end{aligned} \tag{2.18}$$

This completes the proof of part (b).

The critical cone at optimality, from Definition 2.1(d) and (2.8d) and (2.8e), is

$$\begin{aligned} C(x^*, \lambda^*, \psi^*) &= \{ w \in \mathbb{R}^{(2d+1)J} : \nabla c_i(x^*) w = 0, \nabla g_i(x^*) w = 0, 1 \le i \le d \} \\ &= \{ \hat{w} \in \mathbb{R}^{(2d+1)J} : \hat{w}_i = L_{i-1} \hat{w}_{i-1}, 1 \le i \le d \}. \end{aligned} \tag{2.19}$$

We define the vector $u \in \mathbb{R}^{(N+1)J}$ by

$$u_j = \begin{cases} L_j^{(0)} w_1, & 0 \le j \le P_1 \\ L_j^{(P_i-1)} w_{2i} + L_j^{(P_i)} w_{2i+1}, & P_i + 1 \le j \le P_{i+1}, 1 \le i \le d \end{cases}$$

so that for $0 \le i \le d$,

$$\Lambda_i \hat{w}_i = \left[w_{2i+1}^T, u_{P_i+1}^T, \ldots u_{P_{i+1}}^T \right]^T. \tag{2.20}$$

From Definition 2.1(c) the first block row of Λ_i is $[0, I_J]$ for $1 \le i \le d$, and I_J for $i = 0$. Now we consider $w \in C(x^*, \lambda^*, \psi^*)$ and $w \ne \mathbf{0}$. This implies that $w_1 \ne \mathbf{0}$; and since $u_0 = w_1 \ne \mathbf{0}$, we have that $u \ne \mathbf{0}$. Note that since $w \in C(x^*, \lambda^*, \psi^*)$, $L_{P_i}^{(P_i-1)} = \mathbf{0}$, and $L_{P_i}^{(P_i)} = \mathbf{I}_J$, we have from (2.19) that $u_{P_i} = w_{2i+1}$, for $1 \le i \le d$. Substituting this equation in (2.18), using (2.20), using the expression of J_i from Definition 2.1(d), and

using the fact that from Definition 2.1(a) we have that $\nabla_{x_{P_i}}\beta_{P_i} + \nabla_{x_{P_i}}\alpha_{P_i} = \nabla_{x_{P_i}}\theta_{P_i}$ for $1 \le i \le d$, we obtain that

$$\begin{aligned} w^T\nabla_x^2 L(x^*,\lambda^*,\psi^*)w &= u_0^T\nabla_{x_0}\theta_0 u_0 + u_0^T\nabla_{x_1}\theta_0 u_1 \\ &+ \sum_{j=1}^{N-1}\left(u_j^T\nabla_{x_{j-1}}\theta_j u_{j-1} + u_j^T\nabla_{x_j}\theta_j u_j + u_j^T\nabla_{x_{j+1}}\theta_j u_{j+1}\right) \\ &+ u_N^T\nabla_{x_{N-1}}\theta_N u_{N-1} + u_N^T\nabla_{x_N}\theta_N u_N = u^T\left(\nabla^2_{x_{0:N}}\Gamma(x^*_{0:N})\right)u > 0. \end{aligned}$$

This completes the proof of part (c). ∎

3. Stability analysis. The constrained optimization problem (2.8) is now solved with an augmented Lagrangian method. From the Lagrangian function (2.9) and using the notations of (2.8), we define the augmented Lagrangian function.

(3.1)
$$L_A(x,\lambda,\psi,\mu) = \widetilde{\Gamma}(x) - \sum_{i=1}^{d}\lambda_i^T c_i(x) - \sum_{i=1}^{d}\psi_i^T g_i(x) + \frac{\mu}{2}\sum_{i=1}^{d}\left(c_i(x)^T c_i(x) + g_i(x)^T g_i(x)\right)$$

Here $\mu > 0$ is the penalty parameter that helps enforce feasibility. In the rest of this work we assume μ is fixed but large enough so that when λ^* and ψ^* are the Lagrange multipliers of (2.8), the solution x^* of (2.8) is a local minimizer of (3.1). Such a $\mu > 0$ exists from augmented Lagrange theory [18] and Theorem 2.2.

In this section we investigate the condition number of the Hessian matrix for L_A with respect to the number of shooting intervals. In ideal circumstances, the condition number would be bounded above by a constant and thus would prevent exponential growth of the solution in time, which is the signature of instability discussed in §1. Our aim is thus to identify under what circumstances this favorable situation can occur.

For this analysis we use several simplifications to our approach. While our investigations have indicated that similar results can be obtained without making the simplifications, leaving them out would significantly complicate and extend the analysis. We thus keep the number of time points in each shooting interval fixed at k, and we use for all d shooting intervals a fixed time step Δt. Since k is fixed, d grows linearly with N. We consider a constant covariance matrix for model error Q and observation error R for all time steps. The observation mapping is time-dependent linear; that is, $H_i(x_i) = B_i x_i$ for all $0 \le i \le N$ and some $B_i \in \mathbb{R}^{L\times J}$. Note that we allow observation gaps in time, which can be modeled by setting some B_i and the respective observations to $\mathbf{0}$. Theorem 2.2(b), definitions of the constraints (2.8b)–(2.8e) and of the critical cone (2.19), and Definition 2.1(c) imply that the Hessian for L_A at optimality satisfies

$$\begin{aligned} (3.2)\qquad w^T\nabla_x^2 L_A(x^*,\lambda^*,\psi^*,\mu)w &= \sum_{i=0}^{d}\hat{w}_i^T\Lambda_i^T J_i\Lambda_i\hat{w}_i + \mu\sum_{i=1}^{d}\|\hat{w}_i - L_{i-1}\hat{w}_{i-1}\|^2 \\ &+ \left(L_N^{(P_d-1)}w_{2d} + L_N^{(P_d)}w_{2d+1}\right)^T B_N^T R^{-1} B_N\left(L_N^{(P_d-1)}w_{2d} + L_N^{(P_d)}w_{2d+1}\right) \end{aligned}$$

for any $w = (w_1,\ldots,w_{2d+1}) \in \mathbb{R}^{(2d+1)J}$, where we denote $\hat{w}_0 = w_1$, $\hat{w}_i = (w_{2i}, w_{2i+1})$ for $1 \le i \le d$.

We now introduce the definition of the observability matrix for each shooting interval, which is based on the standard one for the linearized system on a given system trajectory [11].

DEFINITION 3.1. *For each* $0 \leq i \leq d$, $P_i \leq j$, *denote* $\prod_{l=P_i}^{j} \nabla M_l(x_l) = \nabla M_j(x_j)\nabla M_{j-1}(x_{j-1})\ldots\nabla M_{P_i}(x_{P_i})$. *Define*

$$C_i^T(x) = \left[B_{P_i}^T, \left(B_{P_i+1}\nabla M_{x_{P_i}}(x_{P_i})\right)^T, \ldots, \left(B_{P_i+k-2}\prod_{l=P_i}^{P_i+k-3}\nabla M_l(x_l)\right)^T \right]$$

as the observability matrix for the (i+1)th shooting interval.

For our work, the importance of the observability condition is that it will ensure that the objective function of (1.3) when applied to the linearized system is positive definite on one shooting interval.

LEMMA 3.2. $C_i(x)$ *being full rank is equivalent to*

$$Q(\boldsymbol{w}) := \sum_{j=P_i}^{P_i+k-2} \left((w_{j+1} - \nabla M_j(x_j)w_j)^T Q^{-1} (w_{j+1} - \nabla M_j(x_j)w_j) + w_j^T B_j^T R^{-1} B_j w_j \right) > 0$$

for any $0 \neq \boldsymbol{w} \in \mathbb{R}^{kJ}$ *and* $0 \leq i \leq d$.

Proof. Suppose there exists $\mathbf{0} \neq s_0 \in \mathbb{R}^J$ such that $C_i s_0 = \mathbf{0}$. Then we define $\mathbf{s} = (s_{P_i}, \ldots, s_{P_i+k-1}) \in \mathbb{R}^{kJ}$ such that $s_{P_i} = s_0$, $s_{P_i+j} = \prod_{l=P_i}^{P_i+j-1} \nabla M_l(x_l) s_0$ for $1 \leq j \leq k-1$. Note that the assumption $C_i s_0 = \mathbf{0}$ and the definition of $\mathbf{s}$ imply that

(3.3)

$$\mathbf{0} = B_{P_i}s_0 = B_{P_i}s_{P_i}, \;\; \mathbf{0} = B_{P_i+j}\prod_{l=P_i}^{P_i+j-1}\nabla M_l(x_l)s_0 = B_{P_i+j}s_{P_i+j}, \; \forall 1 \leq j \leq k-2.$$

Then, (3.3) and the definition of $\mathbf{s}$ give that $Q(\mathbf{s}) = 0$. Note that $\mathbf{s} \neq \mathbf{0}$ since $s_0 \neq \mathbf{0}$.

On the other hand, suppose $Q(\mathbf{s}) = 0$ for some $\mathbf{0} \neq \mathbf{s} = (s_{P_i}, \ldots, s_{P_i+k-1}) \in \mathbb{R}^{kJ}$. Then $B_j s_j = \mathbf{0}$ and $s_{j+1} = \nabla M_j(x_j)s_j$ for $P_i \leq j \leq P_i + k - 2$. Then we have

$$\mathbf{0} = B_{P_i}s_{P_i}, \;\; \mathbf{0} = B_{P_i+j}\prod_{l=P_i}^{P_i+j-1}\nabla M_l(x_l)s_{P_i}, \; \forall 1 \leq j \leq k-2. \tag{3.4}$$

Then, (3.4) implies that $C_i s_{P_i} = \mathbf{0}$. Note that $s_{P_i} \neq \mathbf{0}$ because otherwise $\mathbf{s} = \mathbf{0}$. ∎

A full-rank result holds for the Jacobian matrix of the recursion.

LEMMA 3.3. $\Lambda_i(x_{P_i-1}, x_{P_i})$ *is full rank for* $1 \leq i \leq d$,

Proof. Adapting optimality recursion (2.6) to our simplified model gives

$$\begin{aligned}\widetilde{x}_{P_i+1} = M_{P_i}(x_{P_i}) &+ Q\nabla^{-T}M_{P_i}(x_{P_i})B_{P_i}^T R^{-1}\left(B_{P_i}x_{P_i} - y_{P_i}\right) \\ &+ Q\nabla^{-T}M_{P_i}(x_{P_i})Q^{-1}\left(x_{P_i} - M_{P_i-1}(x_{P_i-1})\right).\end{aligned}$$

and it implies $L_{P_i+1}^{(P_i-1)} = \frac{\partial \widetilde{x}_{P_i+1}}{\partial x_{P_i-1}} = -Q\nabla^{-T}M_{P_i}(x_{P_i})Q^{-1}\nabla M_{P_i-1}(x_{P_i-1})$, which is invertible. Since the first block row of $\Lambda_i(x_{P_i-1}, x_{P_i})$ is $(\mathbf{0}, I)$ and $L_{P_i+1}^{(P_i-1)}$ is the (2,1)th block, $\Lambda_i(x_{P_i-1}, x_{P_i})$ is full rank. ∎

In addition to observability on one shooting interval, we will make slightly stronger assumptions than the ones implied by Lemmas 3.2 and 3.3. That is, we will assume that those bounds hold uniformly with the shooting interval index i.

ASSUMPTION 3.4. *There exist $\gamma_k > 0$ and $\rho_k > 0$ dependent on k but not on i, or d, such that for any $N > 0$, we have the following.*

(a) The observability matrices $C_i(x^*)$ are full rank for $0 \le i \le d$.

(b) Under (a),

$$\sum_{j=P_i}^{P_i+k-2} \left(\left(w_{j+1} - \nabla M_j(x_j^*) w_j\right)^T Q^{-1} \left(w_{j+1} - \nabla M_j(x_j^*) w_j\right) + w_j^T B_j^T R^{-1} B_j w_j \right) \ge \gamma_k \|w\|^2$$

for all $0 \le i \le d$, $w = (w_{P_i}, \ldots, w_{P_i+k-1}) \in \mathbb{R}^{kJ}$.

(c) $\lambda_{min}(\Lambda_i(x^*_{P_i-1}, x^*_{P_i})^T \Lambda_i(x^*_{P_i-1}, x^*_{P_i})) \ge \rho_k$ for all $1 \le i \le d$.

The second set of assumptions characterizes the system, states, and observations as follows.

ASSUMPTION 3.5. *For any $N > 0$,*

(a) $\max_{0\le j\le N} \left(\|x_j^*\|, \|x_B\|\right) \le C_1$ and $\max_{0\le j\le N} \|y_j\| \le C_2$ for some constant $C_1 > 0$ and $C_2 > 0$;

(b) $\max_{0\le j\le N} \|B_j\|_F \le b_0$ for some constant $b_0 > 0$;

(c) $\max_{0\le j\le N} \left(\|\nabla M_j(x_j^*)\|_F, \|\nabla^{-1} M_j(x_j^*)\|_F\right) \le A$ for some constant $A > 0$;

(d) $\max_{0\le j\le N} \left(\|M_j(x_j^*)\|_F\right) \le m_0$ for some constant $m_0 > 0$;

(e) $\max_{0\le j\le N} \|\nabla_{x_j} vec\left(\nabla^T M_j(x_j^*)\right)\|_F \le A_1$ for some constant $A_1 > 0$.

In fact, Assumptions 3.5(d) and (e) are consequences of (a) and the fact that M_j is at least twice continuously differentiable. We nonetheless state them as assumptions so that the bounds we will use in the proof will have convenient references.

We now make a small nonlinearity assumption. It is shown in [1] that for $s \times s$ matrix S and $s \times 1$ vector u and x, we have

$$\nabla_x (Su) = (u^T \otimes I_s) \nabla_x \mathrm{vec}(S) + S \nabla_x u. \tag{3.5}$$

Here we define $M_j^{(2)}(u) := (u^T \otimes I_J) \nabla_{x_j} \mathrm{vec}\left(\nabla^T M_j(x_j^*)\right)$. If u is not a function of x_j, then $M_j^{(2)}(u) = \nabla_{x_j}\left(\nabla^T M_j(x_j^*) u\right)$. Moreover, if the system is linear, then $M_j^{(2)}(u) = \mathbf{0}$; therefore bounds on $M_j^{(2)}(u)$ are bounds limiting nonlinearity. Note that under Assumption 3.5(e), denoting $C_0 = A_1\sqrt{J}$, we have for any $N > 0$ that

$$\max_{0\le j\le N} \|M_j^{(2)}(u)\|_F \le A_1 \|u^T \otimes I_J\|_F \le C_0 \|u\|. \tag{3.6}$$

For our proof, however, we need an even sharper restriction for the nonlinearity described below.

ASSUMPTION 3.6. *There exists $0 \le b_k < \gamma_k$ such that for any $N > 0$,*

$$\max_{0\le j\le N} \|M_j^{(2)}\left(Q^{-1}\left(x_{j+1}^* - M_j(x_j^*)\right)\right)\|_F \le b_k,$$

where γ_k is as defined in Assumption 3.4.

Other than the observability assumption on each shooting interval, Assumptions 3.4 and 3.5 are primarily stating uniformity, and thus are only marginally stronger than the existing assumptions. Assumption 3.6 on the other hand, puts a relatively hard bound on how much nonlinearity we can tolerate in our analysis. At the end of this section we will discuss the effect of this assumption and its significance.

With these definitions and assumptions, we now proceed to the main results of our paper. That is, we now prove that for the nonlinear system satisfying Assumptions 3.4, 3.5, and 3.6, the condition number of the Hessian matrix for the augmented Lagrangian is bounded above. First, we derive a lower bound.

PROPOSITION 3.7. *Under Assumptions 3.4 and 3.6, for any $w \in \mathbb{R}^{kJ}$ and $\|w\| = 1$, we have that $w^T J_i(x^*_{P_i}) w \geq \gamma_k - b_k$ for $0 \leq i \leq d$.*

Proof. Referring back to Definition 2.1(a), we have that

(3.7)

$$\begin{aligned}
\nabla_{x_0}\beta_0 &= \nabla^T M_0(x_0^*)Q^{-1}\nabla M_0(x_0^*) + B_0^T R^{-1} B_0 - M_0^{(2)}\left(Q^{-1}(x_1^* - M_0(x_0^*))\right) + Q_B^{-1}, \\
\nabla_{x_j}\beta_j &= \nabla^T M_j(x_j^*)Q^{-1}\nabla M_j(x_j^*) + B_j^T R^{-1} B_j - M_j^{(2)}\left(Q^{-1}(x_{j+1}^* - M_j(x_j^*))\right), \\
& \qquad\qquad 0 < j \leq N-1 \\
\nabla_{x_j}\alpha_j &= Q^{-1}, \;\; 1 \leq j \leq N, \quad \nabla_{x_{j-1}}\theta_j = -Q^{-1}\nabla^T M_{j-1}(x_{j-1}^*), \;\; 1 \leq j \leq N \\
\nabla_{x_j}\theta_j &= \nabla_{x_j}\alpha_j + \nabla_{x_j}\beta_j, \;\; 0 < j < N \quad \nabla_{x_{j+1}}\theta_j = -\nabla^T M_j(x_j^*)Q^{-1}, \;\; 0 \leq j \leq N-1.
\end{aligned}$$

So for $\|w\| = 1$, referring to Definition 2.1 (d), we have

$$\begin{aligned}
w^T J_i(x^*_{P_i}) w \geq & \sum_{j=P_i}^{P_i+k-2} \left(\left(w_{j+1} - \nabla M_j(x_j^*)w_j\right)^T Q^{-1} \left(w_{j+1} - \nabla M_j(x_j^*)w_j\right) + w_j^T B_j^T R^{-1} B_j w_j \right) \\
& - \sum_{j=P_i}^{P_i+k-2} w_j^T M_j^{(2)}\left(Q^{-1}\left(x_{j+1}^* - M_j(x_j^*)\right)\right) w_j,
\end{aligned}$$

for which equality holds for $1 \leq i \leq d$. For $i = 0$, the difference between the two sides is $w_0^T Q_B^{-1} w_0$, which is non-negative. By Assumption 3.4(b) we have that

$$\sum_{j=P_i}^{P_i+k-2} \left(\left(w_{j+1} - \nabla M_j(x_j^*)w_j\right)^T Q^{-1} \left(w_{j+1} - \nabla M_j(x_j^*)w_j\right) + w_j^T B_j^T R^{-1} B_j w_j \right) \geq \gamma_k,$$

and by Assumption 3.6 we have that $\left|\sum_{j=P_i}^{P_i+k-2} w_j^T M_j^{(2)}\left(Q^{-1}\left(x_{j+1}^* - M_j(x_j^*)\right)\right) w_j\right| \leq b_k$. Thus Proposition 3.7 follows. ∎

We now derive upper bounds in a series of lemmas.

LEMMA 3.8. *Under Assumption 3.5, for each $1 \leq i \leq d$, $P_i + 1 \leq j \leq P_i + k$, and $p = P_i - 1, P_i$, we have that $\|L_j^{(p)}(x^*_{P_i-1}, x^*_{P_i})\|_F \leq C_p^{(j-P_i+1)}$ and $\|L_j^{(0)}(x_0^*)\|_F \leq C_p^j$, where $C_p > 1$ is a constant independent of d.*

Proof. For $0 \leq i \leq d$ and $P_i \leq j \leq P_{i+1} - 1$, define

$$F_{ij} = \nabla M_j(x_j^*) - Q\nabla_{x_j}\left(\nabla^{-T} M_j(x_j^*) B_j^T R^{-1}(y_j - B_j x_j^*)\right),$$

and for $0 \leq i \leq d$ and $P_i + 1 \leq j \leq P_{i+1} - 1$, define

$$\begin{aligned}
G_{ij} &= Q\nabla_{x_j}\left(\nabla^{-T} M_j(x_j^*)Q^{-1}\left(x_j^* - M_{j-1}(x_{j-1}^*)\right)\right) \\
K_{ij} &= -Q\nabla^{-T} M_j(x_j^*)Q^{-1}\nabla M_{j-1}(x_{j-1}^*).
\end{aligned}$$

Also define

$$G_{10} = Q\nabla x_0\left(\nabla^{-T} M_0(x_0^*)Q_B^{-1}(x_0^* - x_B)\right).$$

Then for any $1 \le i \le d$ and $P_i + 1 \le j \le P_i + k$, from optimality recursions (2.7) and the chain rule, the recursion of $L_j^{(P_i)}$ and $L_j^{(P_i-1)}$ can be written as

$$\begin{bmatrix} L_j^{(p)} \\ L_{j-1}^{(p)} \end{bmatrix} = \begin{bmatrix} F_{i,j-1} + G_{i,j-1} & K_{i,j-1} \\ I_J & \mathbf{0} \end{bmatrix} \begin{bmatrix} L_{j-1}^{(p)} \\ L_{j-2}^{(p)} \end{bmatrix}, \tag{3.8}$$

where $p = P_i, P_i - 1$. For the initial shooting interval, the recursion runs through $2 \le j \le P_1$ and $p = 0$. From (2.5), the initialization of the recursion for the initial shooting interval is

$$\begin{bmatrix} L_1^{(0)} \\ L_0^{(0)} \end{bmatrix} = \begin{bmatrix} F_{10} + G_{10} \\ I_J. \end{bmatrix}. \tag{3.9}$$

For the other shooting intervals $1 \le i \le d$, from (2.6), the recursion is initialized by

$$\begin{bmatrix} L_{P_i}^{(P_i-1)} \\ L_{P_i-1}^{(P_i-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ I_J \end{bmatrix}, \qquad \begin{bmatrix} L_{P_i}^{(P_i)} \\ L_{P_i-1}^{(P_i)} \end{bmatrix} = \begin{bmatrix} I_J \\ \mathbf{0} \end{bmatrix}. \tag{3.10}$$

Now we give upper bounds for the propagation matrices. For some $J \times 1$ vector $v(x_j^*)$, by differentiating both sides of $v(x_j^*) = \nabla^T M_j(x_j^*) \nabla^{-T} M_j(x_j^*) v(x_j^*)$ and using equation (3.5), we have that

$$\begin{aligned} \nabla_{x_j} \left(\nabla^{-T} M_j(x_j^*) v(x_j^*) \right) = & -\nabla^{-T} M_j(x_j^*) M_j^{(2)} \left(\nabla^{-T} M_j(x_j^*) v(x_j^*) \right) \\ & + \nabla^{-T} M_j(x_j^*) \nabla v(x_j^*). \end{aligned} \tag{3.11}$$

Now we can give bounds to each part involved in the propagation. By equation (3.11), Assumption 3.5, and equation (3.6), we have that

$$\begin{aligned} &\text{(3.12a)} \quad \| \nabla_{x_j} \left(\nabla^{-T} M_j(x_j^*) B_j^T R^{-1} (y_j - B_j x_j^*) \right) \|_F \\ &\quad \le \| \nabla^{-T} M_j(x_j^*) M_j^{(2)} \left(\nabla^{-T} M_j(x_j^*) B_j^T R^{-1} (y_j - B_j x_j^*) \right) \|_F \\ &\quad + \| \nabla^{-T} M_j(x_j^*) B_j^T R^{-1} B_j \|_F \le C_0 A^2 b_0 \| R^{-1} \|_F (C_2 + b_0 C_1) + A b_0^2 \| R^{-1} \|_F, \\ &\text{(3.12b)} \quad \| \nabla_{x_0} \left(\nabla^{-T} M_0(x_0^*) Q_B^{-1} (x_0^* - x_B) \right) \|_F \\ &\quad \le \| \nabla^{-T} M_0(x_0^*) M_0^{(2)} \left(\nabla^{-T} M_0(x_0^*) Q_B^{-1} (x_0^* - x_B) \right) \|_F \\ &\quad + \| \nabla^{-T} M_0(x_0^*) Q_B^{-1} \|_F \le 2 C_0 A^2 \| Q_B^{-1} \|_F C_1 + A \| Q_B^{-1} \|_F, \\ &\text{(3.12c)} \quad \| \nabla_{x_j} \left(\nabla^{-T} M_j(x_j^*) Q^{-1} \left(x_j^* - M_{j-1}(x_{j-1}^*) \right) \right) \|_F \\ &\quad \le \| \nabla^{-T} M_j(x_j^*) M_j^{(2)} \left(\nabla^{-T} M_j(x_j^*) Q^{-1} \left(x_j^* - M_{j-1}(x_{j-1}^*) \right) \right) \|_F \\ &\quad + \| \nabla^{-T} M_j(x_j^*) Q^{-1} \|_F \le C_0 A^2 \| Q^{-1} \|_F (C_1 + m_0) + A \| Q^{-1} \|_F. \end{aligned}$$

We then have that

$$\begin{aligned} \| F_{ij} \|_F &\overset{\text{(3.12a)}}{\le} A + \| Q \|_F \left(C_0 A^2 b_0 \| R^{-1} \|_F (C_2 + b_0 C_1) + A b_0^2 \| R^{-1} \|_F \right) := F \\ \| G_{ij} \|_F &\overset{\text{(3.12c)}}{\le} \| Q \|_F \left(C_0 A^2 \| Q^{-1} \|_F (C_1 + m_0) + A \| Q^{-1} \|_F \right) := G_1 \\ \| K_{ij} \|_F &\le A^2 \| Q \|_F \| Q^{-1} \|_F := K \\ \| G_{10} \|_F &\overset{\text{(3.12b)}}{\le} \| Q \|_F \left(2 C_0 A^2 \| Q_B^{-1} \|_F C_1 + A \| Q_B^{-1} \|_F \right) := G_0. \end{aligned}$$

Let $G = \max(G_1, G_0)$. Then, bounding each term in the propagation relations (3.8), (3.9), and (3.10) by its Forbenius norm, we have for $1 \le i \le d$, $P_i + 1 \le j \le P_i + k$, and $p = P_i - 1, P_i$ that

$$\begin{aligned}\|L_j^{(p)}\|_F &\le \left\| \begin{bmatrix} L_j^{(p)} \\ L_{j-1}^{(p)} \end{bmatrix} \right\|_F \le \left\| \begin{bmatrix} F_{i,j-1} + G_{i,j-1} & K_{i,j-1} \\ I_J & \mathbf{0} \end{bmatrix} \right\|_F \left\| \begin{bmatrix} L_{j-1}^{(p)} \\ L_{j-2}^{(p)} \end{bmatrix} \right\|_F \\ &\le \left(\sqrt{J + K^2 + (F+G)^2}\right)^{j-P_i} \sqrt{J} \le \left(\sqrt{J + K^2 + (F+G)^2}\right)^{j-P_i+1} := C_p^{j-P_i+1},\end{aligned}$$

and for the initial shooting interval, similarly we have for $1 \le j \le P_1$ that

$$\|L_j^{(0)}\|_F \le \left(\sqrt{J + K^2 + (F+G)^2}\right)^{j-1} \sqrt{J + (F+G)^2} \le \left(\sqrt{J + K^2 + (F+G)^2}\right)^{j} = C_p^j.$$

□

LEMMA 3.9. *Under Assumptions 3.5 and 3.6 and using notations in Definition 2.1(d), for each $1 \le i \le d$ we have that $\|J_0(x_0^*)\|_F, \|J_i(x_{P_i-1}^*, x_{P_i}^*)\|_F \le C_J$ for some $C_J > 0$ independent of d.*

Proof. Because of the block tridiagonal structure of J_i for $0 \le i \le d$, we have that

$$\begin{aligned}\|J_i\|_F &\le \sum_{j=P_i+1}^{P_{i+1}-1} \left(\|\nabla_{x_{j-1}}\theta_j\|_F + \|\nabla_{x_j}\theta_j\|_F + \|\nabla_{x_{j+1}}\theta_j\|_F\right) \\ &\quad + \|\nabla_{x_{P_i}}\beta_{P_i}\|_F + \|\nabla_{x_{P_i+1}}\theta_{P_i}\|_F + \|\nabla_{x_{P_{i+1}-1}}\theta_{P_{i+1}}\|_F + \|\nabla_{x_{P_{i+1}}}\alpha_{P_{i+1}}\|_F \\ &\overset{(3.7)}{\le} \sum_{j=P_i+1}^{P_{i+1}-1} \left(2A\|Q^{-1}\|_F + \|Q^{-1}\|_F + A^2\|Q^{-1}\|_F + b_0^2\|R^{-1}\|_F + b_k\right) \\ &\quad + 2A\|Q^{-1}\|_F + \|Q^{-1}\|_F + A^2\|Q^{-1}\|_F + b_0^2\|R^{-1}\|_F + b_k \\ &\le k\left(2A\|Q^{-1}\|_F + \|Q^{-1}\|_F + A^2\|Q^{-1}\|_F + b_0^2\|R^{-1}\|_F + b_k\right) \\ &:= C_J.\end{aligned}$$

□

PROPOSITION 3.10. *For any $w \in \mathbb{R}^{(2d+1)J}$ and $\|w\| = 1$, we have that $w^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) w \le U_k$ for some $U_k > 0$ independent of d.*

Proof. For $0 \le i \le d$, using Lemmas 3.8 and 3.9 and referring to Definition 2.1, we have that $\|J_i\|_F \|\Lambda_i\|_F^2 \le 2(k+1) C_J C_p^{2k}$. Then, from (3.2), it follows that

$$\begin{aligned}w^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) w &= \sum_{i=0}^{d} \hat{w}_i^T \Lambda_i^T J_i \Lambda_i \hat{w}_i + \mu \sum_{i=1}^{d} \|\hat{w}_i - L_{i-1}\hat{w}_{i-1}\|^2 \\ &+ \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right)^T B_N^T R^{-1} B_N \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right) \\ &\le 2(k+1) C_J C_p^{2k} + b_0^2 \|R^{-1}\|_F C_p^{2k} + \mu(1 + 2C_p^k)^2.\end{aligned}$$

Defining U_k to be the last quantity above completes the proof. □

We are now in a position to state and prove our main result.

THEOREM 3.11. *Under Assumptions 3.4, 3.5, and 3.6, the condition number of the Hessian matrix for the augmented Lagrangian is bounded above independent of the number of shooting intervals, d. That is,*

$$\kappa\left(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)\right) \le \frac{U_k}{(\gamma_k - b_k)\min(\rho_k, 1)}.$$

Proof. For any $w \in \mathbb{R}^{(2d+1)J}$ and $\|w\| = 1$, using Proposition 3.7 and Assumption 3.4(c), we have that

$$\begin{aligned} & w^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) w = \sum_{i=0}^{d} \hat{w}_i^T \Lambda_i^T J_i \Lambda_i \hat{w}_i + \mu \sum_{i=1}^{d} \|\hat{w}_i - L_{i-1}\hat{w}_{i-1}\|^2 \\ & + \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right)^T B_N^T R^{-1} B_N \left(L_N^{(P_d-1)} w_{2d} + L_N^{(P_d)} w_{2d+1}\right) \\ \geq \quad & \sum_{i=0}^{d} \hat{w}_i^T \Lambda_i^T J_i \Lambda_i \hat{w}_i \geq (\gamma_k - b_k) \sum_{i=0}^{d} \|\Lambda_i \hat{w}_i\|^2 \\ \overset{\text{Assumption 3.4}}{\geq} \quad & (\gamma_k - b_k) \left(\rho_k \sum_{i=1}^{d} \|\hat{w}_i\|^2 + \|\hat{w}_0\|^2 \right) \geq (\gamma_k - b_k) \min(\rho_k, 1). \end{aligned}$$

Combining with Proposition 3.10, we obtain

$$\kappa\big(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)\big) = \frac{\lambda_{max}\left(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)\right)}{\lambda_{min}\left(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)\right)} \leq \frac{U_k}{(\gamma_k - b_k)\min(\rho_k, 1)},$$

which completes the proof. ∎

Discussion. An interpretation of Theorem 3.11 is that, under observability Assumption 3.4 and small nonlinearity Assumption 3.6, the condition number of the multiple shooting problem is bounded above with the number of multiple shooting intervals d. This prevents the exponential increase of the solution, which we define as instability, and thus makes the multiple shooting problem computable. We note that the upper bounds of the lemmas preceding Theorem 3.11 allow for exponential increase *within the shooting interval*; but as long as observability holds, this increase stops at the end of a shooting interval. As for Assumptions 3.6, we note that the amount of nonlinearity needs to be upper bounded by the lower bound γ_k that is related to observability by Lemma 3.2. This points out that the bound on nonlinearity in Assumption 3.6 is not absolute; it only needs to be small compared with how much information can be found in the observations. That is, increasing the measurement space would increase the lower eigenvalue of $\sum B_i^T R^{-1} B_i$ and thus γ_k, which in turn would increase the prospects for Assumption 3.6 to hold.

Another important question is whether these assumptions are necessary. While an if and only if statement between observability and the bounded condition number of the multiple shooting Lagrangian probably does not hold, some of the assumptions are necessary in the following way. As we can see from Appendix A, multiple shooting without observations still results in exponential increase of the condition number and thus of the solution. Therefore some amount of observability, or, otherwise said, state space coverage by data, is necessary. As we can see from Appendix B, without multiple shooting the condition number of the Hessian matrix for the single shooting function (1.7) also increases exponentially and thus is unstable. We conclude that some form of observability and multiple shooting are necessary to obtain a stability result as Theorem 3.11.

4. Recursive gradient evaluation. When implementing minimization of the augmented Lagrangian function (3.1), gradient evaluation is required. In this section, we describe a recursive method for computing the gradient of (3.1) that fits into our memory-saving framework.

First we derive the gradients of the augmented Lagrangian function. Note that $\theta_j(\widetilde{x}_{j-1}, \widetilde{x}_j, \widetilde{x}_{j+1}) = \mathbf{0}$ for all $P_i + 1 \le j \le P_{i+1} - 1$, $0 \le i \le d$, and $\theta_0(x_0, \widetilde{x}_1) = \mathbf{0}$. For the first interval we obtain that

$$\begin{aligned} \nabla_{x_0} L_A(x,\lambda,\psi,\mu) &= L_{P_1}^{(0)^T}\left(\nabla_{x_{P_1}}\phi_{P_1-1}(\widetilde{x}_{P_1-1},\widetilde{x}_{P_1}) + \lambda_1 - \mu c_1(x)\right) \\ &+ L_{P_1-1}^{(0)^T}(\psi_1 - \mu g_1(x)) \quad + L_{P_1}^{(0)^T}\cancel{\left(\sum_{j=1}^{P_1-1}\theta_j(\widetilde{x}_{j-1},\widetilde{x}_j,\widetilde{x}_{j+1})\right)} + \cancel{\theta_0(x_0,\widetilde{x}_1)}. \end{aligned} \tag{4.1}$$

For $1 \le i \le d-1$, we obtain that

$$\begin{aligned} \nabla_{x_{P_i-1}} L_A(x,\lambda,\psi,\mu) &= L_{P_{i+1}}^{(P_i-1)^T}\left(\cancel{\sum_{j=P_i+1}^{P_{i+1}-1}\theta_j(\widetilde{x}_{j-1},\widetilde{x}_j,\widetilde{x}_{j+1})} + \nabla_{x_{P_{i+1}}}\phi_{P_{i+1}-1}(\widetilde{x}_{P_{i+1}-1},\widetilde{x}_{P_{i+1}})\right) \\ &+ L_{P_{i+1}}^{(P_i-1)^T}(\lambda_{i+1} - \mu c_{i+1}(x)) + L_{P_{i+1}-1}^{(P_i-1)^T}(\psi_{i+1} - \mu g_{i+1}(x)) + \mu g_i(x) - \psi_i. \\ \nabla_{x_{P_i}} L_A(x,\lambda,\psi,\mu) &= L_{P_{i+1}}^{(P_i)^T}\left(\cancel{\sum_{j=P_i+1}^{P_{i+1}-1}\theta_j(\widetilde{x}_{j-1},\widetilde{x}_j,\widetilde{x}_{j+1})} + \nabla_{x_{P_{i+1}}}\phi_{P_{i+1}-1}(\widetilde{x}_{P_{i+1}-1},\widetilde{x}_{P_{i+1}})\right) \\ &+ L_{P_{i+1}}^{(P_i)^T}(\lambda_{i+1} - \mu c_{i+1}(x)) + L_{P_{i+1}-1}^{(P_i)^T}(\psi_{i+1} - \mu g_{i+1}(x)) + \mu c_i(x) - \lambda_i + \beta_{P_i}(x_{P_i},\widetilde{x}_{P_i+1}). \end{aligned}$$

For the last shooting interval, we obtain that

$$\begin{aligned} \nabla_{x_{P_d-1}} L_A(x,\lambda,\psi,\mu) &= L_N^{(P_d-1)^T}\left(\cancel{\sum_{j=P_d+1}^{N-1}\theta_j(\widetilde{x}_{j-1},\widetilde{x}_j,\widetilde{x}_{j+1})} + \theta_N(\widetilde{x}_{N-1},\widetilde{x}_N)\right) \\ &- \psi_d + \mu g_d(x), \\ \nabla_{x_{P_d}} L_A(x,\lambda,\psi,\mu) &= L_N^{(P_d)^T}\left(\cancel{\sum_{j=P_d+1}^{N-1}\theta_j(\widetilde{x}_{j-1},\widetilde{x}_j,\widetilde{x}_{j+1})} + \theta_N(\widetilde{x}_{N-1},\widetilde{x}_N)\right) \\ &+ \beta_{P_d}(x_{P_d},\widetilde{x}_{P_d+1}) - \lambda_d. \end{aligned}$$

Note that the derivatives are composed of a matrix-vector product for which the vector can be computed through one forward recursion similar to the one for the states. The Jacobian matrix $L_{P_{i+1}}^{(P_i)}$, however, needs to also be computed by forward recursion, and it turns out to be dense. The computation thus would require $O(J^2)$ storage and inhibit the low-memory advantage of our approach. Instead, we compute the matrix-vector product using a backward recursion separately on each multiple shooting interval, as follows. Since the evaluation procedure is the same for each interval, we illustrate our method with the first interval (assuming it has length N').

The target of our algorithm is to compute $v^T L_{N'}^{(0)}$ for some constant vector v. This algorithm can then be used to compute the gradient components defined in the beginning of this section. For example, for computing the first component (4.1) we note that we have two such matrix-vector products, where N' is, succesively P_1 and $P_1 - 1$ and v is succesively $\left(\nabla_{x_{P_1}}\phi_{P_1-1}(\widetilde{x}_{P_1-1},\widetilde{x}_{P_1}) + \lambda_1 - \mu c_1(x)\right)$ and $(\psi_1 - \mu g_1(x))$. Similar embeddings hold for all other gradient components.

The computation of $v^T L_{N'}^{(0)}$ proceeds as follows. The optimality recursion states that $\theta_j(\widetilde{x}_{j-1}(x_0), \widetilde{x}_j(x_0), \widetilde{x}_{j+1}(x_0)) = 0$ for $1 \le j \le N' - 1$. Differentiating with

respect to x_0 gives

$$L_{j+1}^{(0)} = -(\nabla_{x_{j+1}}\theta_j)^{-1}\left((\nabla_{x_{j-1}}\theta_j)L_{j-1}^{(0)} + (\nabla_{x_j}\theta_j)L_j^{(0)}\right). \tag{4.2}$$

Now we write the recursion ansatz and substitute (4.2) to obtain

$$v^T L_{N'-l+1}^{(0)} = c_l^T L_{N'-l}^{(0)} + b_l^T L_{N'-l-1}^{(0)} = -c_l^T(\nabla_{x_{N'-l}}\theta_{N'-l-1})^{-1}(\nabla_{x_{N'-l-2}}\theta_{N'-l-1})L_{N'-l-2}^{(0)}$$
$$+\left(b_l^T - c_l^T(\nabla_{x_{N'-l}}\theta_{N'-l-1})^{-1}(\nabla_{x_{N'-l-1}}\theta_{N'-l-1})\right)L_{N'-l-1}^{(0)} := c_{l+1}^T L_{N'-l-1}^{(0)} + b_{l+1}^T L_{N'-l-2}^{(0)}$$

for $1 \le l \le N'-2$, where c_{l+1} and b_{l+1} for $l = 2, \ldots, N'-2$ are defined by sought-after recursions

$$c_{l+1}^T = b_l^T - c_l^T(\nabla_{x_{N'-l}}\theta_{N'-l-1})^{-1}(\nabla_{x_{N'-l-1}}\theta_{N'-l-1}), \tag{4.3}$$

$$b_{l+1}^T = -c_l^T(\nabla_{x_{N'-l}}\theta_{N'-l-1})^{-1}(\nabla_{x_{N'-l-2}}\theta_{N'-l-1}). \tag{4.4}$$

Then the matrix-vector product of interest can be expressed as $v^T L_{N'}^{(0)} = c_{N'-1}^T L_1^{(0)} + b_{N'-1}^T L_0^{(0)}$, where $c_{N'-1}$ and $b_{N'-1}$ are obtained through recursions (4.3) and (4.4). It is a backward recursion with respect to the usage of state information x_j. The initial values for the recursion are

$$c_1^T = -v^T(\nabla_{x'_N}\theta_{N'-1})^{-1}(\nabla_{x_{N'-1}}\theta_{N'-1}), \quad b_1^T = -v^T(\nabla_{x'_N}\theta_{N'-1})^{-1}(\nabla_{x_{N'-2}}\theta_{N'-1}),$$

obtained by total differentiation of $\theta_{N'-1}(\widetilde{x}_{N'-2}(x_0), \widetilde{x}_{N'-1}(x_0), \widetilde{x}_{N'}(x_0)) = 0$.

Since the recursion can be computed separately on each shooting interval, the total storage does not exceed the number of multiple shooting checkpoints plus the length of an interval, which adds up to $2d+1+N/(d+1)$. We can use checkpointing within the shooting interval to reduce the storage even further, but we do not pursue that avenue here.

5. Numerical results. In this section, we apply our multiple shooting method to Burgers' equation in order to verify some of our theoretical findings. This is a one-spatial-dimension, time-dependent, partial differential equation that exhibits both diffusion and nonlinear advection. Since implementation of new ideas in an operational environment is a development-intensive process, in many research references discussing new state estimation methods Burgers' equation is considered an important first test of a method [2, 13, 14, 26].

The partial differential equation describing it is the following:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial(u^2)}{\partial x} = \nu\frac{\partial^2 u}{\partial x^2}; \quad u(0,t) = u(1,t) = 0; \quad u(x,0) = u_0(x), \tag{5.1}$$

where $\nu > 0$ is viscosity coefficient and $(x,t) \in (0,1) \times (0,T)$.

We denote by u_j^m the unknown value at grid coordinates $(j\Delta x, m\Delta t)$ and $\Delta x = 1/J$. We use a centered finite-difference discretization [2]:

$$\frac{u_j^{m+1} - u_j^m}{\Delta t} + \frac{(u_{j+1}^m)^2 - (u_{j-1}^m)^2}{4\Delta x} - \frac{\nu}{(\Delta x)^2}(u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1}) = 0. \tag{5.2}$$

To demonstrate the benefits of multiple shooting, we choose parameters for which the single shooting method in [1] exhibits instability. To make the problem closer to

intended application target, we also experiment with larger model error and sparser observations, which are known to be more difficult [1]. We compare the solution of the multiple shooting method with that obtained from directly minimizing the full-memory function (1.3) in our examples. Note that the full-memory problem itself is not without difficulties: it cannot be solveed to high accuracy by LBFGS in any of our examples within 2,000 iterations. The norm of gradient of (1.3) decreases slowly approaching the end and never gets below 10^{-6}. In this section, we refer to the approach of minimizing the full-memory function as 4D-Var for brevity, although our example is (1+1)D.

5.1. Results for Burgers' equation. We choose $\Delta x = 1/500$, $\Delta t = \Delta x/500$, background state $x_B = \sin(\pi x)$, and background covariance $Q_B = 0.01I$. We generate the initial state u_0 by sampling from the background distribution, namely, $u_0 \sim \mathcal{N}(x_B, Q_B)$. The rest of the states are generated by model propagation plus a model error term, namely, $u_{t+1} = M_t(u_t) + \eta_t$ for $0 \le t < N$, where $\eta_t \sim \mathcal{N}(0, Q)$ and $Q = (\Delta t)^2 \mathrm{diag}(2, 1, \ldots, 1, 2)$ is the covariance of model error. The observations are generated by applying $H_t(u_t) = \sin(u_t)$ to the underlying states $U = \{u_0, \ldots, u_N\}$ plus a mean zero normal observation error term to mimic the action of a noisy nonlinear operator. We note that for analytic simplicity our theoretical results consider only the linear observation operator case; but we expect the nonlinear one to be even harder, so we could use the results to validate the outcome of multiple shooting. The covariance of observation error is chosen as $R = 0.01I$. The observations are made with a gap of 10 steps in time and space.

Our aim is to minimize the augmented Lagrangian function (3.1). For achieving our limited-memory purpose, we use LBFGS [18] with $p = 6$ stored vectors. To obtain an initial point for minimization, we first perturb the underlying state U by the error of the background distribution. This mimics the situation where the estimation does not start cold; in other words, initial estimates of the states do exist from previous runs of the algorithm. On each shooting interval, we run the 4D-Var minimization of (2.10) with LBFGS for 200 iterations to get a "warm start" state $\{w_0, \ldots, w_N\}$. Note that 4DVar is run only in the beginning on each interval separately on which p trajectories are stored. The largest amount of memory required is then $\max\{2d + 1 + \frac{N}{d+1}, (p+1)\frac{N}{d+1}\}$ state vectors. We also note that applying LBFGS to the 4DVar problem on the entire horizon requires $(p+1)N$ state vectors, which is most times d times larger. We add that in this and in the other numerical sections, it proved difficult to find another starting strategy that will reliably produce a point from which the multiple shooting algorithm will converge. On the other hand, this strategy does work and does not alter the storage reduction benefits of our approach.

The checkpoints of the warm start state $\{w_0, w_{P_1-1}, w_{P_1}, \ldots, w_{P_d-1}, w_{P_d}\}$ are then used as the initial point for minimizing (3.1). The Lagrangian multiplier and penalty parameters are initially chosen as $\lambda_i^{(0)} = \mathbf{0}$, $\psi_i^{(0)} = \mathbf{0}$, $\mu^{(0)} = 10$ and are subject to the usual Lagrange multiplier updates [18].

In Table 5.1 we tabulate the number of checkpoint pairs d, number of stored vectors, and percentage of storage over full-memory storage for each of the examples in this section. For $N = 800$, $d = 12$ is the smallest number of checkpoint pairs to make the computation stable. For each $800 \le N \le 1600$, the corresponding d is chosen so that $d/\sqrt{N} = 12/\sqrt{800}$. For $N = 2400$, d is chosen to satisfy $d/N = 12/800$. We choose $d \propto N$ for $N = 800$ and 2400 to demonstrate that the method is stable for increasing N and hence to verify Theorem 3.11. For $1000 \le N \le 1600$, we choose another relation $d \propto \sqrt{N}$ to demonstrate empirically the consequences of a more

N	800	1000	1200	1400	1600	2400
d	12	14	15	17	19	36
storage	434	469	525	546	560	455
$\frac{\text{storage}}{(p+1)N}$	7.8%	6.7%	6.3%	5.6%	5.0%	2.7%

Table 5.1: Number of checkpoint pairs d and maximal storage for $\Delta t = \Delta x/500$.

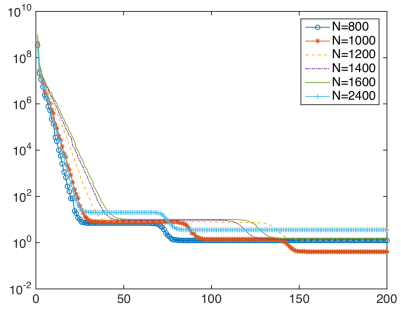

Fig. 5.1: Function value of (3.1) at each iteration of LBFGS for $\Delta t = \Delta x/500$ and $N = 800, 1000, 1200, 1400, 1600, 2400$.

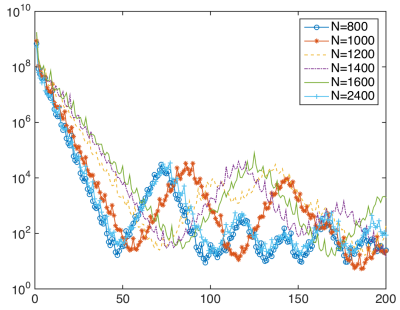

Fig. 5.2: Gradient norm of (3.1) at each iteration of LBFGS for $\Delta t = \Delta x/500$ and $N = 800, 1000, 1200, 1400, 1600, 2400$.

aggressive checkpointing schedule.

Figure 5.1 compares the function value reduction of (3.1) at each iteration of LBFGS for increasing time horizon. For $800 \leq N \leq 1600$, the rate of the initial descent (before iteration 50) becomes smaller as N increases, which indicates slower convergence for increasing N. This means that a more aggressive checkpoint schedule (e.g., $d \propto \sqrt{N}$) can lead to slower convergence. In contrast, the rate of descent for $N = 2400$ is closer to that of $N = 800$ and much larger than those of $1000 \leq N \leq 1600$. It indicates that the method not only is stable but converges with similar speed for increasing N if d is allowed to increase linearly in N. Figure 5.2 shows the norm of gradient at each iteration. Figure 5.3 shows the Frobenius norm of constraints c_i, g_i, $1 \leq i \leq d$ at each iteration. Figure 5.4 plots the Euclidean distance scaled by Δx of each iteration to the checkpoints of the full-memory 4D-Var solution. Note that the distance is not scaled by the number of states and is expected to increase with d.

In this experiment, we see significant reduction (by 8–9 orders of magnitude) for both the function value and the norm of gradient, even if the gradient did not decrease to a point that triggered the Lagrange multiplier update. Figure 5.5 plots the solution surface of multiple shooting and 4D-Var when $N = 2400$. Both of them approach a perturbed version of the noise-free solution. Figure 5.6 compares multiple shooting and 4D-Var solutions at fixed time and space nodes. Note that the two solutions are both close to the underlying state so that the trajectories overlap for most of the part. Although the problem is not solved to high accuracy as suggested by the norm of the gradient and norm of the constraint, we conclude that it does approach the 4D-Var solution.

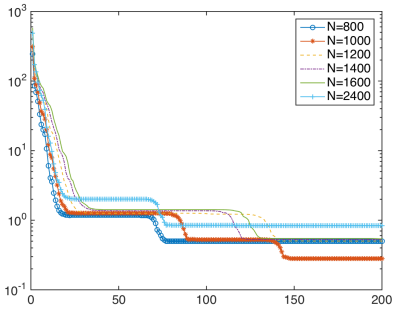

Fig. 5.3: Norm of constraint at each iteration of LBFGS in minimizing (3.1) for $\Delta t = \Delta x/500$ and $N = 800, 1000, 1200, 1400, 1600, 2400$.

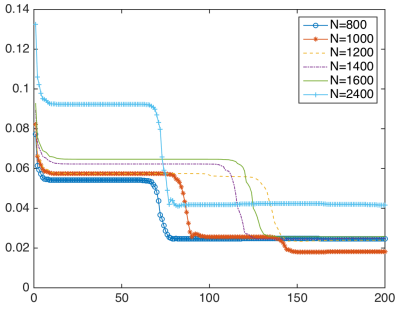

Fig. 5.4: Distance to 4D-Var solution at each iteration of LBFGS in minimizing (3.1) for $\Delta t = \Delta x/500$ and $N = 800, 1000, 1200, 1400, 1600, 2400$.

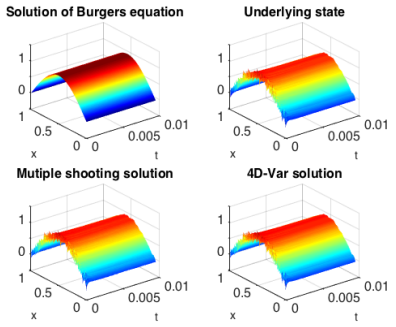

Fig. 5.5: Exact solution of Burgers equation (top left), underlying state (top right) and states estimated with multiple shooting and 4D-Var for $\Delta t = \Delta x/500$ and $N = 2400$.

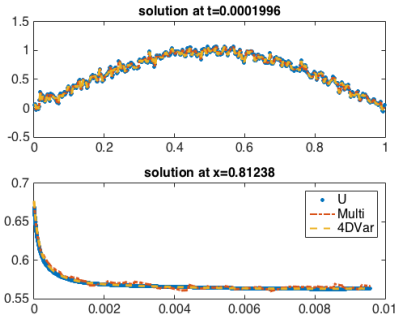

Fig. 5.6: Underlying state, multiple shooting solution and 4D-Var solution at fixed time and space node for $\Delta t = \Delta x/500$ and $N = 2400$.

From the simulations we see that keeping N/d fixed (at its lowest value) results in faster convergence compared with the alternatives. We thus conclude that the statement of Theorem 3.11 is satisfied, although its conditions are stronger than the case tested here (we did not enforce small nonlinearity and linearity of the observation operator). However, for the case of smaller N (e.g., 800 to 1600), even increasing d slower than linear in N (e.g., $\sqrt{N}$) would give stable results and thus even more memory savings at a cost of somewhat slower convergence.

5.2. Larger model error. In this section, we experiment with increased model error. We choose $\Delta x = 1/500$, $\Delta t = \Delta x/1000$, and a background covariance matrix $Q_B = 0.01I$. The covariance for the model error and observation error are chosen to be $10^{-3}I$. Observations are reduced to every 10 steps in time and every 100 steps in space. To initialize the minimization of (3.1), we run the 4D-Var minimization on one interval, and for the next interval we run 4D-Var constrained at the checkpoint

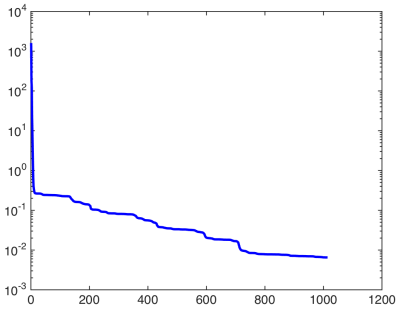

Fig. 5.7: Function value of (3.1) at each iteration of LBFGS for $\Delta t = \Delta x/1000$, $N = 500$ and $d = 38$.

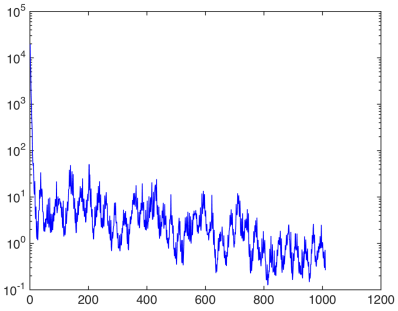

Fig. 5.8: Gradient norm of (3.1) at each iteration of LBFGS for $\Delta t = \Delta x/1000$, $N = 500$ and $d = 38$.

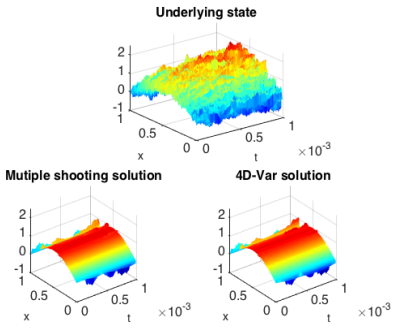

Fig. 5.9: Underlying states, solution surface of multiples shooting and 4D-Var for $\Delta t = \Delta x/1000$, $N = 500$ and $d = 38$.

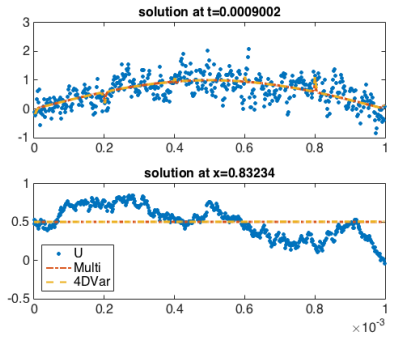

Fig. 5.10: Underlying state, multiple shooting solution and 4D-Var solution at fixed time and space nodes for $\Delta t = \Delta x/1000$, $N = 500$ and $d = 38$.

by the solution from the previous interval.

Figure 5.7 shows the augmented Lagrangian function value decrease for $N = 500$ and number of checkpoint pairs $d = 38$. Figure 5.8 shows the norm of the gradient. Figure 5.9 compares the full-memory 4D-Var solution with that of multiple shooting. Increased model error results in the rough surface of the underlying states plot in Figure 5.9. Figure 5.10 compares the 4DVar and multiple shooting solutions at fixed time and space nodes. Note that the two solutions are close to each other so that their trajectories overlap.

Both the function value and the norm of the gradient converge slower after some significant initial progress. Since the norm of the gradient stalls and fails to progress below 0.1, we do not observe either Lagrangian multiplier or penalty parameter update during the experiments. However, both the function value and the norm of the gradient achieve 4 to 6 orders of magnitude decrease, and the multiple shooting solution approaches reasonably well the full-memory 4D-Var solution. Clearly the problem has too much noise for the estimates to be close to the underlying state. However, the

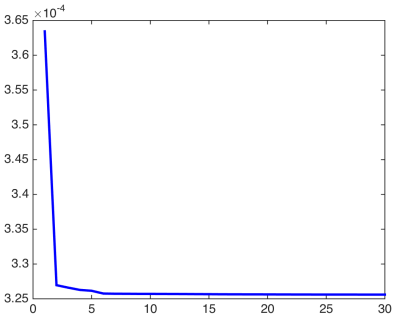

Fig. 5.11: Function value of (3.1) at each iteration of LBFGS for $N = 300$ and $d = 30$.

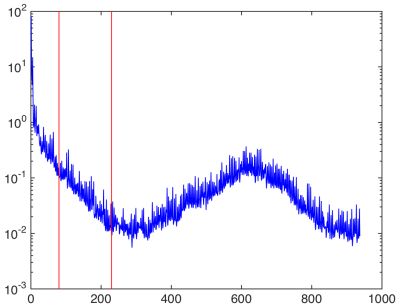

Fig. 5.12: Gradient norm of (3.1) at each iteration of LBFGS for $N = 300$ and $d = 30$. Reference line indicates Lagrangian multiplier updates.

approach does show that multiple shooting has a performance comparable to that of 4DVar, with much less memory, and that is the goal of this paper.

With the same parameters as those in [1, §5.2.5] but with a much longer horizon, $N = 500$ as opposed to $N = 110$, our method is able to produce iterations of moderate size, make nontrivial progress through minimization, and result in solutions comparable to that of full-memory method for a longer time horizon. Counting the storage during warm start, gradient evaluation, and stored vectors of LBFGS, the maximal number of states stored at any time of the algorithm is 91 and is about 18.2% of the total number of states N. The storage used by multiple shooting is 2.6% of the memory used by full-memory minimization using LBFGS with 6 vectors.

5.3. Sparser observations. In this section, we consider the case where observations are made sparser in both time and space. We choose $\Delta x = 1/700$, $\Delta t = \Delta x/34$, and background covariance as $Q_B = 10^{-3}I$. The covariance matrix for the model error and observation error are $Q = 10^{-8}I$ and $0.01I$, respectively. Observations are made every 30 steps in time and every 200 steps in space. The initial point for multiple shooting is the same warm-start point described in the precedent section. The parameters are the same as those in [1, §5.2.5] but with a longer horizon. We take $N = 300$ as opposed to $N = 32$ in [1], and we take number of checkpoint pairs $d = 30$.

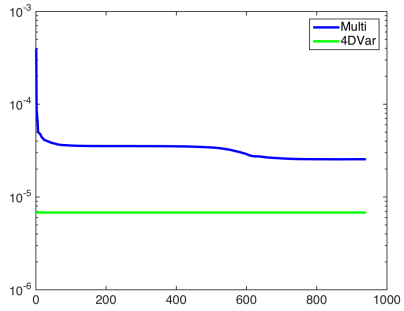

Fig. 5.13: Norm of constraint at each iteration of LBFGS in minimizing (3.1) for $N = 300$ and $d = 30$.

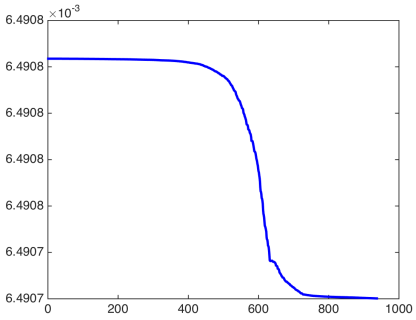

Fig. 5.14: Distance to 4D-Var solution at each iteration of LBFGS in minimizing (3.1) for $N = 300$ and $d = 30$.

We note that this setup is significantly far from satisfying the observability condition. Indeed, the rank of the observability matrix in Definition 3.1 cannot be larger than 8, whereas Theorem 3.11 required a full rank, that is, 701.

For this experiment, Figure 5.11 shows the decrease of function value (3.1). Only the first 30 iterations are plotted since the function value stalls afterward. Lagrangian multipliers are updated at iteration 80 and 230, as shown by the vertical reference line in Figure 5.12. Figure 5.13 shows the norm of constraints $c_i(x)$ and $g_i(x)$ at each iteration. The horizontal reference line plotted is the norm of constraint for the 4D-Var solution. Figure 5.14 shows the Euclidean distance of each iteration to the 4D-Var solution scaled by Δx. The decrease in the norm of the gradient is significant (3–4 orders of magnitude), and the norm of the constraint is reduced by about 1 order of magnitude. The distance to the 4D-Var solution shows little progress compared with the initial guess obtained by running 4D-Var on each shooting interval, but Figures 5.13 and 5.14 suggest the reason is primarily that our warm-starting using 4D-Var on each shooting interval produces an initial point for multiple shooting close to the 4D-Var solution itself. On the other hand, even if in the distance to the 4D-Var solution there is not much progress beyond the warm start, the gradient is significantly reduced, and we can evaluate the convergence properties of the method, running LBGFS to detect whether we see an improvement, while needing less memory than 4D-Var with LBFGS (only 3.4% of the latter's). Therefore the multiple shooting method provides an improvement over 4D-Var with LBFGS in terms of memory and over single shooting in terms of stability even in this case, which is significantly outside the applicability of Theorem 3.11.

6. Conclusions. Determining the best state estimation for dynamical systems with model error raises new challenges in developing algorithms that reduce storage while maintaining stability. The reason is that, as opposed to the strongly constrained setups where only the initial state is free, all the states of a trajectory contribute to the number of degrees of freedom.

We present an approach where the number of degrees of freedom is reduced by the optimality conditions, as we previously introduced in [1], but now coupled with a multiple shooting approach in an augmented Lagrangian framework to improve stability. The multiple shooting approach can use a reverse recursion scheme on each

shooting interval to ensure that the memory requirements for computing one gradient of the augmented Lagrangian never exceed $2d+1+\frac{N}{d+1}$ state vectors, where $d+1$ is the number of shooting intervals and N is the length of the horizon. The full-memory data assimilation method, on the other hand, needs to store $N+1$ state vectors when evaluating its gradient. We prove in Theorem 3.11 that under an observability assumption and when the nonlinearity is small relative to the parameter characterizing the observability, the condition number of the augmented Lagrangian matrix is bounded above, irrespective of the number of shooting intervals. This ensures that the multiple shooting approach is stable: the method does not exhibit exponentially increasing error for an increasing size of the assimilation interval. This is a feature shared by neither the single shooting approach derived from [1] nor by the multiple shooting approach without observations. Therefore both multiple shooting and sufficiently informative observations appear to be necessary for stability to occur.

Our numerical simulations on cases described in [1] validate these points. First, for all of them the single shooting method showed an exponential increase of the solution and ran into overflow. For both small model error and larger model error setups, the multiple shooting approach converges to a solution close to that of the full-memory method while using only a fraction of the memory needed by the latter, never more than 8%. To achieve convergence, we needed to use the full-memory approach but only on the smaller, shooting intervals to create a good initial point for our multiple shooting approach. In the case of sparse observations, this initialization strategy was responsible for much of the improvement of the method in terms of distance to the full-memory 4D-Var solution, while using only 3.4% of the memory of the latter. But with that initialization strategy, which does not alter our maximum memory count, we reliably obtained reductions in the augmented Lagrangian gradients and solutions close to the ones of the full-memory approach. We are not aware of another optimization-based approach to reduce the memory requirements of weakly constrained data assimilation approaches. From the numerical experiments and the theory, we conclude that, particularly in the data-rich case, the multiple shooting method appears promising at reducing memory and producing a point of a quality comparable to that of the full-memory case without the instability of the previous single shooting approach.

We plan to explore new initialization strategies that empirically appear to be important for the robustness of the overall method. The method also has good potential for paralellism, although in that case the memory saving is less of a benefit. An interesting question would be to tie the stability of multiple shooting to a condition requiring enough information in the observations but weaker than observability on one shooting interval. We have observed the good behavior of the multiple shooting aproach in several such instances, but it is unclear how such a condition might be expressive enough and practical.

Acknowledgment. This material is based upon work supported by the U.S. Department of Energy, Office of Science, under Contract No. DE-AC02-06CH11357. We are grateful to our colleagues Xiaoyan Zeng and Emil Constantinescu for feedback and assistance at the beginning of this project.

REFERENCES

[1] Mihai Anitescu, Xiaoyan Zeng, and Emil M Constantinescu, *A low-memory approach for best-state estimation of hidden markov models with model error*, SIAM Journal on

Numerical Analysis, 52 (2014), pp. 468–495.
[2] A. Apte, D. Auroux, and M. Ramaswamy, *Variational data assimilation for discrete Burgers equation*, in Electronic Journal of Differential Equations, vol. 19, 2010, pp. 15–30.
[3] Thomas Bengtsson, Peter Bickel, Bo Li, et al., *Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems*, in Probability and statistics: Essays in honor of David A. Freedman, Institute of Mathematical Statistics, 2008, pp. 316–334.
[4] Roger Daley, *Atmospheric data analysis*, no. 2, Cambridge university press, 1993.
[5] DP Dee, *Testing the perfect-model assumption in variational data assimilation*, in Proc. Second Int. Symp. on Assimilation of Observations in Meteorology and Oceanography, Citeseer, 1995, pp. 225–228.
[6] FRANÇOIS-XAVIER LE DIMET and Olivier Talagrand, *Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects*, Tellus A, 38 (1986), pp. 97–110.
[7] A. Doucet and A.M. Johansen, *A Tutorial on Particle Filtering and Smoothing: Fifteen years later*, The Oxford Handbook of Nonlinear Filtering, Oxford University Press., (2009).
[8] James Glimm, Shuling Hou, Yoon-Ha Lee, David H Sharp, and Kenny Ye, *Sources of uncertainty and error in the simulation of flow in porous media*, Computational & Applied Mathematics, 23 (2004), pp. 109–120.
[9] Andreas Griewank, *Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation*, Optimization Methods and software, 1 (1992), pp. 35–54.
[10] Andrew H Jazwinski, *Stochastic processes and filtering theory*, Courier Corporation, 2007.
[11] Rudolf Kalman, *On the general theory of control systems*, IRE Transactions on Automatic Control, 4 (1959), pp. 110–110.
[12] Rudolph Emil Kalman, *A new approach to linear filtering and prediction problems*, Journal of Fluids Engineering, 82 (1960), pp. 35–45.
[13] Eugenia Kalnay, *Atmospheric modeling, data assimilation, and predictability*, Cambridge university press, 2003.
[14] John M Lewis, Sivaramakrishnan Lakshmivarahan, and Sudarshan Dhall, *Dynamic data assimilation: a least squares approach*, vol. 13, Cambridge University Press, 2006.
[15] Magnus Lindskog, Dick Dee, Yannick Tremolet, Erik Andersson, Gabor Radnoti, and Mike Fisher, *A weak-constraint four-dimensional variational analysis system in the stratosphere*, Quarterly Journal of the Royal Meteorological Society, 135 (2009), pp. 695–706.
[16] MJ Martin, MJ Bell, and Nancy K Nichols, *Estimation of systematic error in an equatorial ocean model using data assimilation*, International Journal for Numerical Methods in Fluids, 40 (2002), pp. 435–444.
[17] IM Navon, *Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography*, Dynamics of Atmospheres and Oceans, 27 (1998), pp. 55–79.
[18] J. Nocedal and S.J. Wright, *Numerical optimization*, Springer Verlag, 1999.
[19] David Orrell, L Smith, J Barkmeijer, and TN Palmer, *Model error in weather forecasting*, Nonlinear processes in geophysics, 8 (1999), pp. 357–371.
[20] TN Palmer, GJ Shutts, R Hagedorn, FJ Doblas-Reyes, Thomas Jung, and M Leutbecher, *Representing model uncertainty in weather and climate prediction*, Annu. Rev. Earth Planet. Sci., 33 (2005), pp. 163–193.
[21] L.R. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*, Proceedings of the IEEE, 77 (1989), pp. 257–286.
[22] L. Rabiner and B. Juang, *An introduction to hidden Markov models*, IEEE ASSp Magazine, 3 (1986), pp. 4–16.
[23] Dan Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*, John Wiley & Sons, 2006.
[24] Yannick Trémolet, *Accounting for an imperfect model in 4d-var*, Quarterly Journal of the Royal Meteorological Society, 132 (2006), pp. 2483–2504.
[25] ———, *Model-error estimation in 4d-var*, Quarterly Journal of the Royal Meteorological Society, 133 (2007), pp. 1267–1280.
[26] Francesco Uboldi and Masafumi Kamachi, *Time-space weak-constraint data assimilation for nonlinear models*, Tellus A, 52 (2000), pp. 412–421.
[27] Milija Zupanski, Dusanka Zupanski, Tomislava Vukicevic, Kenneth Eis, and Thomas Vonder Haar, *Cira/csu four-dimensional variational data assimilation system*, Monthly Weather Review, 133 (2005), pp. 829–843.

Appendix A. Multiple shooting with zero observability. In this section, we prove that for a class of linear systems, under zero observability, the condition number of the Hessian matrix of augmented Lagrangian has an exponential lower bound. Hence the multiple shooting method is not stable if there are no observations.

We consider the model propagation mapping to be time independent, that is, $M(x_j) = Ax_j$, and $B_j = \mathbf{0}$ for $0 \le j \le N$. We assume A has at least one real eigenvalue with modulus strictly larger than 1. With this model specification, J_i are identical for all $1 \le i \le d$ and so are Λ_i. For simplicity, we denote them respectively as J_1 and Λ_1 for $1 \le i \le d$. The expanded forms of J_1 and Λ_1 are

$$J_1 = \begin{bmatrix} A^TQ^{-1}A & -A^TQ^{-1} & & & \mathbf{0} \\ -Q^{-1}A & A^TQ^{-1}A + Q^{-1} & \ddots & & \\ & \ddots & \ddots & & \\ & & \ddots & A^TQ^{-1}A + Q^{-1} & -A^TQ^{-1} \\ \mathbf{0} & & & -Q^{-1}A & Q^{-1} \end{bmatrix},$$

$$\Lambda_1 = \begin{bmatrix} \mathbf{0} & I \\ -QA^{-T}Q^{-1}A & A + QA^{-T}Q^{-1} \\ \vdots & \vdots \\ L_{P_i+1}^{(P_i-1)}(x_{P_i-1}, x_{P_i}) & L_{P_i+1}^{(P_i)}(x_{P_i-1}, x_{P_i}) \end{bmatrix}.$$

For $p = P_i, P_i - 1$, adapting the optimality recursions (2.7) to the linear system under consideration and applying the chain rule, we have that the recursion of $L_{P_i+j}^{(p)}$ for $0 \le j \le k-1$ is

$$L_{P_i+j+1}^{(p)} = (A + QA^{-T}Q^{-1})L_{P_i+j}^{(p)} - QA^{-T}Q^{-1}AL_{P_i+j-1}^{(p)}. \tag{A.1}$$

Denote L_1 to be the last two block rows of Λ_1.

LEMMA A.1. *Denote* $\hat{\Lambda} = \Lambda_1 \begin{bmatrix} I \\ A \end{bmatrix}$. *Then,* $J_1\hat{\Lambda} = \mathbf{0}$.

Proof. We first prove that for $1 \le j \le k$, the jth block of $\hat{\Lambda}$ is $(\hat{\Lambda})_j = A^j$ by induction. It is evident for $j = 1, 2$. Suppose it is true for all $j \le j_0$, $2 \le j_0 \le k-1$. Then by recursion (A.1),

$$\begin{aligned} (\hat{\Lambda})_{j_0+1} &= L_{P_i+j_0}^{(P_i-1)} + L_{P_i+j_0}^{(P_i)}A \\ &= (A + QA^{-T}Q^{-1})(L_{P_i+j_0-1}^{(P_i-1)} + L_{P_i+j_0-1}^{(P_i)}A) \\ &\quad - (QA^{-T}Q^{-1}A)(L_{P_i+j_0-2}^{(P_i-1)} + L_{P_i+j_0-2}^{(P_i)}A) \\ &= (A + QA^{-T}Q^{-1})(\hat{\Lambda})_{j_0} - (QA^{-T}Q^{-1}A)(\hat{\Lambda})_{j_0-1} \\ &= A^{j_0+1}. \end{aligned}$$

A direct multiplication completes the proof. ∎

PROPOSITION A.2. *Let* $|\lambda| > 1$, $\lambda \in \mathbb{R}$ *be an eigenvalue of* A. *Denote* $\lambda_k = \lambda^{k-1}$. *Then, for the linear system under consideration, we have that*

$$\kappa\big(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)\big) \ge \frac{\lambda_{min}(Q_B^{-1})}{\mu}|\lambda_k|^{2(d-1)}.$$

Proof. For any $s = (s_1, \ldots, s_{2d+1}) \in \mathbb{R}^{(2d+1)J}$, denote $\hat{s}_0 = s_1$, $\hat{s}_i = (s_{2i}, s_{2i+1})$ for $1 \le i \le d$. Then from Theorem 2.2 (b) we have that

$$s^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) s = s_1^T \Lambda_0^T J_0 \Lambda_0 s_1 + \sum_{i=1}^{d} \hat{s}_i^T \Lambda_1^T J_1 \Lambda_1 \hat{s}_i + \mu \|\hat{s}_1 - L_0 s_1\|^2 + \mu \sum_{i=2}^{d} \|\hat{s}_i - L_1 \hat{s}_{i-1}\|^2. \tag{A.2}$$

Consider $s = (s_1, \ldots, s_{2d+1}) \in \mathbb{R}^{(2d+1)J}$ such that $s_1 = 0$, $s_{2i} = \lambda_k^{i-1} s_2$, $s_{2i+1} = \lambda s_{2i}$ for $1 \le i \le d$, and let $\|s_2\| = 1$ be the eigenvector of A corresponding to λ, that is, $As_2 = \lambda s_2$. Then $\hat{s}_i = \begin{bmatrix} I \\ A \end{bmatrix} s_{2i}$ for $1 \le i \le d$, which gives that

$$\hat{s}_i^T \Lambda_1^T J_1 \Lambda_1 \hat{s}_i = s_{2i}^T \hat{\Lambda}^T J_1 \hat{\Lambda} s_{2i} = 0, \tag{A.3}$$

where the last equality follows from Lemma A.1.

Since L_1 consists of the last two block rows of Λ_1, we have that $L_1 \begin{bmatrix} I \\ A \end{bmatrix} = \begin{bmatrix} A^{k-1} \\ A^k \end{bmatrix}$. Hence by the definition of s for $2 \le i \le d$, we obtain that

$$\hat{s}_i - L_1 \hat{s}_{i-1} = \begin{bmatrix} I \\ A \end{bmatrix} s_{2i} - \begin{bmatrix} A^{k-1} \\ A^k \end{bmatrix} s_{2(i-1)} = \mathbf{0}. \tag{A.4}$$

Using (A.3) and (A.4) in (A.2), we obtain that

$$s^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) s = \mu \|\hat{s}_1\|^2 \le \mu(\|s_2\|^2 + |\lambda|^2 \|s_2\|^2) = \mu(1 + |\lambda|^2).$$

From the definition of s, we have that

$$\|s\|^2 = \sum_{i=1}^{d} \|s_{2i}\|^2 + \|s_{2i+1}\|^2 = (1 + |\lambda|^2) \sum_{i=1}^{d} |\lambda_k|^{2(i-1)} \ge (1 + |\lambda|^2) |\lambda_k|^{2(d-1)}.$$

Hence we have that

$$\lambda_{min}(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)) \le \frac{s^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) s}{\|s\|^2} \le \mu |\lambda_k|^{-2(d-1)}. \tag{A.5}$$

On the other hand, let $t = (t_1, \ldots, t_{2d+1}) \in \mathbb{R}^{(2d+1)J}$ be such that $\|t_1\| = 1$ and $t_i = 0$ for all $2 \le i \le 2d+1$. J_0 differs from J_1 by only the (1,1)th block element so that $(J_0)_{(1,1)} = (J_1)_{(1,1)} + Q_B^{-1}$. Then

$$t^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) t = t_1^T \Lambda_0^T J_0 \Lambda_0 t_1 + \mu t_1^T L_0^T L_0 t_1 \ge t_1^T \Lambda_0^T J_0 \Lambda_0 t_1 \ge \lambda_{min}(Q_B^{-1}).$$

Hence we have that

$$\lambda_{max}(\nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu)) \ge \frac{t^T \nabla_x^2 L_A(x^*, \lambda^*, \psi^*, \mu) t}{\|t\|^2} \ge \lambda_{min}(Q_B^{-1}). \tag{A.6}$$

Combining equation (A.5) and (A.6) completes the proof. ∎

Appendix B. Single shooting condition number. In this section, we prove that for a certain class of linear systems that satisfy the observability condition, the condition number of the Hessian matrix for the single shooting function (1.7) has an exponential lower bound in N. Hence the single shooting method is not stable for this class of systems.

We consider linear time-independent systems such that $M(x_i) = Ax_i$ and $H(x_i) = Bx_i$. Denote $C_1 = QA^{-T}Q^{-1} + A + QA^{-T}B^TR^{-1}B$ and $C_2 = QA^{-T}Q_B^{-1} + A + QA^{-T}B^TR^{-1}B$. We have the following.

PROPOSITION B.1. *For linear systems satisfying*

(a) $C_1C_2 - I = C_2^2$,

(b) there exist eigenvalues λ_1 *and* λ_2 *of* C_2 *such that* $|\lambda_1| > 1$ *and* $|\lambda_1| > |\lambda_2| \neq 0$,

(c) $QA^{-T}Q^{-1}A = I_J$.

We have

$$\kappa\Big(\nabla^2_{x_0}\hat{\Gamma}(x_0^*)\Big) \geq \begin{cases} \frac{C}{N}\left|\frac{\lambda_1}{\lambda_2}\right|^{2(N-1)}, & |\lambda_2| \geq 1 \\ \frac{C}{N}|\lambda_1|^{2(N-1)}, & |\lambda_2| < 1 \end{cases}$$

for some constant $C > 0$, *where* x_0^* *is the first component of a local minimizer of* $\Gamma(x_{0:N})$ (1.3).

Note: At the end of this section, we give an example of a linear system satisfying conditions (a)–(c) with observation matrix B being full rank, namely, with full observability.

Proof. It is shown in [1, Theorem 3] that x_0^* is a local minimizer of $\hat{\Gamma}(x_0)$ and that

$$\nabla_{x_0}\hat{\Gamma}(x_0^*) = \theta_0(x_0^*, \lambda_1) + \sum_{j=1}^{N-1} {L_j^{(0)}}^T \theta_j(\lambda_{j-1}, \lambda_j, \lambda_{j+1}) + {L_N^{(0)}}^T \theta_N(\lambda_{N-1}, \lambda_N), \tag{B.1}$$

where $L_j^{(0)}$, $0 \leq j \leq N$ are as defined in Definition 2.1(b).

Applying the chain rule and the optimality conditions (1.4), (1.5), and (1.6) to (B.1), we obtain that the Hessian matrix for the single shooting function (1.7) is

$$\nabla^2_{x_0}\hat{\Gamma}(x_0^*) = \Lambda_s^T J_s \Lambda_s, \tag{B.2}$$

where Λ_s is $(N+1)J \times J$ dimensional and J_s is $(N+1)J \times (N+1)J$ dimensional. They are defined as

$$\Lambda_s^T = \begin{bmatrix} {L_0^{(0)}}^T & {L_1^{(0)}}^T & \dots & {L_N^{(0)}}^T \end{bmatrix}$$

$$J_s = \begin{bmatrix} Q_B^{-1} + B^TR^{-1}B + A^TQ^{-1}A & -A^TQ^{-1} & & & \mathbf{0} \\ -Q^{-1}A & C_3 + A^TQ^{-1}A & \ddots & & \\ & \ddots & \ddots & & \\ & & \ddots & C_3 + A^TQ^{-1}A & -A^TQ^{-1} \\ \mathbf{0} & & & -Q^{-1}A & C_3 \end{bmatrix},$$

where $C_3 := Q^{-1} + B^TR^{-1}B$. Denote $d_j(x) = \begin{bmatrix} L_j^{(0)}x \\ L_{j-1}^{(0)}x \end{bmatrix}$, for $1 \leq j \leq N$. Then, for

$1 \leq j \leq N-1$, from the recursion for the derivatives (3.8) and (3.9), we have that

$$\begin{aligned} d_{j+1}(x) &= \begin{bmatrix} QA^{-T}Q^{-1} + A + QA^{-T}B^TR^{-1}B & -QA^{-T}Q^{-1}A \\ I_J & \mathbf{0} \end{bmatrix} d_j(x) \\ &= \begin{bmatrix} QA^{-T}Q^{-1} + A + QA^{-T}B^TR^{-1}B & -I_J \\ I_J & \mathbf{0} \end{bmatrix} d_j(x) \\ &= \begin{bmatrix} C_1 & -I_J \\ I_J & \mathbf{0} \end{bmatrix} d_j(x) := Dd_j(x), \end{aligned}$$

and

$$\begin{aligned} d_1(x) &= \begin{bmatrix} QA^{-T}B^TR^{-1}B + A + QA^{-T}Q_B^{-1} \\ I_J \end{bmatrix} x \\ &= \begin{bmatrix} C_2 \\ I_J \end{bmatrix} x := \hat{C}_2 x. \end{aligned}$$

For any eigenvector v of C_2 with corresponding eigenvalue λ, we have from condition (a) that $Dd_1(v) = \lambda d_1(v)$. Hence for $1 \leq j \leq N$, we have that $d_j(v) = \lambda^{j-1} d_1(v)$. Denoting $\widetilde{Q} = (I, -A)^T Q^{-1}(I, -A)$ and using (B.2), we have that

$$\begin{aligned} v^*\nabla^2_{x_0}\hat{\Gamma}(x_0^*)v &= v^*\Lambda_s^T J_s \Lambda_s v \\ &= v^*Q_B^{-1}v + \sum_{j=0}^{N}(L_j^{(0)}v)^*B^TR^{-1}B(L_j^{(0)}v) \\ &\quad + \sum_{j=0}^{N-1}(L_{j+1}^{(0)}v - AL_j^{(0)}v)^*Q^{-1}(L_{j+1}^{(0)}v - AL_j^{(0)}v) \\ &= v^*Q_B^{-1}v + \sum_{j=0}^{N}(L_j^{(0)}v)^*B^TR^{-1}B(L_j^{(0)}v) + \sum_{j=1}^{N} d_j(v)^*\widetilde{Q}d_j(v) \\ &= v^*Q_B^{-1}v + \sum_{j=0}^{N}(L_j^{(0)}v)^*B^TR^{-1}B(L_j^{(0)}v) + v^*\hat{C}_2^T\widetilde{Q}\hat{C}_2 v \sum_{j=1}^{N}|\lambda|^{2(j-1)}, \end{aligned}$$

where $\hat{C}_2^T\widetilde{Q}\hat{C}_2 = (Q_B^{-1} + B^TR^{-1}B)^TA^{-1}QA^{-T}(Q_B^{-1} + B^TR^{-1}B)$ is positive definite.

Let v_1 and v_2 be eigenvectors of C_2 corresponding respectively to λ_1 and λ_2 as defined in condition (b), and $\|v_1\| = 1$, $\|v_2\| = 1$. Then we have

$$\begin{aligned} \lambda_{max}(\nabla^2_{x_0}\hat{\Gamma}(x_0^*)) &\geq \frac{v_1^*\nabla^2_{x_0}\hat{\Gamma}(x_0^*)v_1}{\|v_1\|^2} \\ &\geq v_1^*\hat{C}_2^T\widetilde{Q}\hat{C}_2 v_1 \sum_{j=1}^{N}|\lambda_1|^{2(j-1)} \\ &\geq \lambda_{min}(\hat{C}_2^T\widetilde{Q}\hat{C}_2)|\lambda_1|^{2(N-1)} \end{aligned} \tag{B.3}$$

and

$$\begin{aligned}\lambda_{min}(\nabla^2_{x_0}\hat{\Gamma}(x_0^*)) &\leq \frac{v_2^*\nabla^2_{x_0}\hat{\Gamma}(x_0^*)v_2}{\|v_2\|^2}\\ &\leq \lambda_{max}(Q_B^{-1}) + \lambda_{max}(B^TR^{-1}B)\|d_1(v_2)\|^2\sum_{j=1}^{N}|\lambda_2|^{2(j-1)}\\ &\quad + \lambda_{max}(B^TR^{-1}B) + v_2^*\hat{C}_2^T\widetilde{Q}\hat{C}_2v_2\sum_{j=1}^{N}|\lambda_2|^{2(j-1)}\\ &\leq 2U + 2U\sum_{j=1}^{N}|\lambda_2|^{2(j-1)}\\ &\leq \begin{cases}4UN|\lambda_2|^{2(N-1)}, & |\lambda_2|\geq 1\\ 4UN, & |\lambda_2|<1\end{cases},\end{aligned} \tag{B.4}$$

where $U = \max\left(\lambda_{max}(Q_B^{-1}), \lambda_{max}(B^TR^{-1}B)\lambda_{max}(\hat{C}_2^T\hat{C}_2), \lambda_{max}(B^TR^{-1}B), \lambda_{max}(\hat{C}_2^T\widetilde{Q}\hat{C}_2)\right)$.

Equations (B.3) and (B.4) give that

$$\kappa\left(\nabla^2_{x_0}\hat{\Gamma}(x_0^*)\right) \geq \begin{cases}\frac{\lambda_{min}(\hat{C}_2^T\widetilde{Q}\hat{C}_2)}{4UN}\left|\frac{\lambda_1}{\lambda_2}\right|^{2(N-1)}, & |\lambda_2|\geq 1\\ \frac{\lambda_{min}(\hat{C}_2^T\widetilde{Q}\hat{C}_2)}{4UN}|\lambda_1|^{2(N-1)}, & |\lambda_2|<1\end{cases}.$$

Letting $C = \frac{\lambda_{min}(\hat{C}_2^T\widetilde{Q}\hat{C}_2)}{4U}$ completes the proof. ∎

Consider an example for which $Q = A = diag(2, 1, \ldots, 1)$ for all $0 \leq j \leq N-1$, $Q_B = diag(4, \frac{3}{2}, \ldots, \frac{3}{2})$, and $B^TR^{-1}B = diag(\frac{7}{4}, \frac{4}{3}, \ldots, \frac{4}{3})$ such that B is full rank. Then, $C_1 = diag(\frac{17}{4}, \frac{10}{3}, \ldots, \frac{10}{3})$ and $C_2 = diag(4, 3, \ldots, 3)$ so that all three conditions in Proposition B.1 are satisfied. For this example, even with full observability, single shooting is not stable.