OPERATOR SPLITTING METHODS IN 4D VARIATIONAL DATA ASSIMILATION

LIJIAN JIANG*, CRAIG C. DOUGLAS†, AND YALCHIN EFENDIEV‡

Abstract. In this paper, we give a framework 4D variational data assimilation (4D-Var) in Hilbert Spaces. In functional partial differential equation setting, the adjoint equation method and Marchuk-Strang operator splitting are discussed. Convergence analysis of the operator splitting method is investigated for the functional 4D-Var.

Key words. 4D-Var, Cost functional, Marchuk-Strang operator splitting method

AMS subject classifications. 65J15, 65K10, 49K27

1. Introduction. There is a steadily growing interest in variational data assimilation [1, 11, 16]. Data assimilation systems use two sources of data: observations and a recent forecast (or "background") valid at a known time. In many real life problems, observation sets are distributed in three dimensional space plus time, corresponding to four dimensional variational data assimilation (4D-Var). 4D-Var is a method of estimating a set of parameters by optimizing the fit between the solution of a model and a set of observations which the model is meant to predict. The unknown model parameters may be the model's initial conditions or boundary conditions. The problem of determining the model parameters is very important and complex and has become a science in itself. The goal of 4D-Var is to incorporate actual observations (e.g., satellite, radar, ship, land surface and so on) into mathematical and numerical models in order to create a unified, complete description of some substance (e.g., atmosphere or chemical pollutant) arising from nature. This can be used by scientific communities to study important phenomena associated with the substance. 4D-Var is a process where a state forecast and observations are combined to produce a best (optimal) estimate or an analysis of the state [11].

Typically, optimization of parameters in 4D-Var can be regarded as a class of inverse problems and the parameters are often constrained by evolution differential operators. So 4D-Var can be reduced to a PDE-constrained optimization problem. Without loss of generality, we will use a parameter which represents the initial condition of a nonlinear evolution differential operator. Therefore the 4D-Var problem (inverse problem) can be posted as follows: "what initial condition will seed the model to best predict the known observations?" A practical implementation of the optimization process requires a fast and accurate evaluation of the gradient of a optimization functional which may be provided by an adjoint model. Generally the adjoint model relies on the existence of the forward model itself which is run many times in an iterative procedure.

In this paper, we provide a framework for 4D-Var in functional setting and investigate the symmetric operator splitting methods in the 4D-Var problem.

*Department of Mathematics, Texas A & M University, College Station, TX 77843-3368. email: ljjiang@math.tamu.edu

†Departments of Computer Science and Mechanical Engineering, University of Kentucky, Lexington, KY 40506-0046; Department of Computer Science, Yale University, New Haven, CT. email: craig.douglas@uky.edu

‡Department of Mathematics, Texas A & M University, College Station, TX 77843-3368. email: efendiev@math.tamu.edu

The paper is organized as follows. In Section 2, we introduce a framework for 4D-Var in functional setting, where the observation operator and constrain differential operator are both nonlinear in Hilbert spaces. In section 3, we carry out the Marchuk-Strang multicomponent operator splitting to the functional 4D-Var and make convergence analysis. In Section 4, we draw some conclusions.

2. 4D Variational Data Assimilation in Functional Setting.

2.1. A Framework for 4D-Var. In this section, we discuss a general framework for 4D-Var. 4D-Var is a generalization of 3D-Var for observations that are distributed in time. We set the framework in a nonlinear functional defined in Hilbert spaces.

Let V be Hilbert space with inner product $(\cdot,\cdot)_V$ and $U = \{\phi : [t_0, T] \longrightarrow V\}$ be a Hilbert space. In the paper, we always assume that $(\cdot,\cdot) = (\cdot,\cdot)_V$ unless otherwise stated. Let u^B represent the background of the initial values (the initial guess in the assimilation) and B be an associated linear covariance operator of the estimated background error. Let H be an observation operator which depends on time t (without specification, we always assume that $H = H(t)$ in the paper). Basically, H maps space U onto a proper subspace of U. Let u_{obs} be the real observations depending on time t. The linear covariance operator $R^{-1} = R^{-1}(t)$ accounts for observations and representativeness errors. In its general form a cost function can be defined as

$$J(u^0, u) = \frac{1}{2}(u^0 - u^B, B^{-1}(u^0 - u^B)) + \frac{1}{2}(Hu - u_{obs}, R^{-1}(Hu - u_{obs}))_U, \quad (2.1)$$

where $u \in U$, $u^0 \in V$ and

$$(\cdot,\cdot)_U = \int_{t_0}^{T} (\cdot,\cdot)(s)ds.$$

The cost functional measures the difference between the model output u and the observation u_{obs} and the deviation of the solution from the background state u^B.

REMARK 2.1. *In many real life situations, the observation u_{obs} is not defined continuously in a time interval $[t_0, T]$, but is evaluated at a set of discrete time moments $\{t_k\}_{k=0}^N$. For these cases, the cost functional is redefined as*

$$J(u^0, u) = \frac{1}{2}(u^0 - u^B, B^{-1}(u^0 - u^B)) + \frac{1}{2}\sum_{k=0}^{N}(H_k u^k - u^k_{obs}, R_k^{-1}(H_k u^k - u^k_{obs})). \quad (2.2)$$

REMARK 2.2. *If the observation u_{obs} is not defined continuously in domain Ω, but is evaluated at a set of discrete locations $\{x_l\}_{l=0}^L$, then the cost functional is redefined as*

$$J(u^0, u) = \frac{1}{2}(u^0 - u^B, B^{-1}(u^0 - u^B)) + \frac{1}{2}(Hu - u_{obs}, R^{-1}(Hu - u_{obs})\delta(x - x_l))_U, \quad (2.3)$$

where $\delta(\cdot)$ is the Dirac function.

REMARK 2.3. *$J(u^0, u)$ is called sensitivity functional in the adjoint optimization.*

Let $S(t)$ is a predefined forecast model (a nonlinear operator) on $[t_0, t]$. Generally, 4D variational assimilation is defined as the minimization problem,

$$\hat{u^0} = Arg \min\{J(u^0, u) : u(t) = S(t)u^0\},$$

where the model states $u(t)$ are subject to the forward model equation,

$$u(t) = S(t)u^0. \tag{2.4}$$

Let $\bar{S}(t)$ be the linearization of $S(t)$. Corresponding to (2.4), a tangent linear model is defined by

$$\delta u(t) = \bar{S}(t)\delta u^0. \tag{2.5}$$

Let $\bar{S}^*(t)$ be the adjoint of $\bar{S}(t)$, $\bar{H}^*(t)$ the adjoint of $\bar{H}(t)$, and $\bar{H}(t)$ a linearization of $H(t)$. Let ∇_{u^0} be the Fréchet derivative operator (or gradient operator) with respect to u^0 in the space U and $\nabla^2_{u^0}$ the second derivative operator with respect to u^0. Since $u(t) = S(t)u^0$, a direct calculation gives rise to

$$\nabla_{u^0} J = B^{-1}(u^0 - u^B) + \int_{t_0}^{T} \bar{S}^* \bar{H}^* R^{-1}(Hu - u_{obs})(s)ds \tag{2.6}$$

and

$$\nabla^2_{u^0} J(u^0) = B^{-1} + \int_{t_0}^{T} \bar{S}^* \bar{H}^* R^{-1} \bar{H}\bar{S} ds. \tag{2.7}$$

For numerical approximation of (2.1) we need to discretize time. Without loss of generality, we can discuss the cost functional defined in (2.2) instead of the time discretization of (2.1) or (2.3). This is because the cost functional is the same as in (2.2) in many real life problems and the time discretization of (2.1) is similar to (2.2). For example, a time discretization approximation of (2.1) can be written for a time step length of $\tau = \frac{T-t_0}{N}$ as

$$J(u^0, u) \approx \frac{1}{2}(u^0 - u^B, B^{-1}(u^0 - u^B)) + \frac{1}{2}\sum_{k=1}^{N} \tau(H_k u^k - u^k_{obs}, R_k^{-1}(H_k u^k - u^k_{obs})).$$

In this paper, we will use the cost functional in (2.1) to discuss the time continuous 4D-Var and use the functional in (2.2) to discuss the time discrete (numerical) 4D-Var.

Hence, 4D-Var is just a nonlinear constrained optimization problem that is very difficult to solve in the general case. For numerical approximation, 4D-Var can be simplified with two hypothesis:

Hypothesis 1: *The forecast model can be expressed as the product of intermediate forecast steps. Let $S_{[t,t+\tau]}$ be the forecast step from t to $t+\tau$, i.e., $u(t+\tau) = S_{[t,t+\tau]}u(t)$. Consequently, for $t_k = t_0 + k\tau$,*

$$u(t_k) = S_{[t_{k-1},t_k]} \cdots S_{[t_1,t_2]} S_{[t_0,t_1]} u^0.$$

Remark 2.4. *Hypothesis 1 usually means that the forecast model is the integration of a numerical prediction model starting with u^0 as the initial condition. Hypothesis 1 is suitable for an operator splitting method for the PDE problem.*

Hypothesis 2: *HSu^0 has a first order Taylor expansion around u^B at any observation time, i.e.,*

$$HSu^0 = HSu^B + \bar{H}\bar{S}(u^0 - u^B),$$

where we recall that $\bar{H} = \bar{H}(t)$ is the linearization of the observation operator $H(t)$ and $\bar{S} = \bar{S}(t)$ is the tangent linear model of $S(t)$, i.e., the differential (perturbation) of $S(t)$.

REMARK 2.5. *Hypothesis 2 is a tangent linear hypothesis and implies that*

$$\nabla_{u^0} HS(u^0) = \bar{H}\bar{S}.$$

Let $\bar{S}_{[t_{k-1},t_k]} = \nabla_u S_{[t_{k-1},t_k]} u|_{t=t_{k-1}}$, $\bar{S}_{[t_0,t_k]} = \Pi_{i=1}^{k} \bar{S}_{[t_{i-1},t_i]}$, and $\bar{H}_k = \nabla_u Hu|_{t=t_k}$. Consequently, it can be verified that

$$\nabla_{u^0} J = \nabla_{u^0} J = B^{-1}(u^0 - u^B) + \sum_{k=0}^{N} \bar{S}^*_{[t_k,t_0]} \bar{H}^*_k R_k^{-1}(H_k u^k - u^k_{obs}) \tag{2.8}$$

and

$$\nabla^2_{u^0} J = B^{-1} + \sum_{k=0}^{N} \bar{S}^*_{[t_k,t_0]} \bar{H}^*_k R_k^{-1} \bar{H}_K \bar{S}_{[t_0,t_k]}. \tag{2.9}$$

2.2. Constraint Specified by A Nonlinear Evolution Operator. We now discuss the case that the constraint $u(t) = S(t)u_0$ is specified by a nonlinear evolution operator A defined on a Hilbert space U. The variational data assimilation problem associated with A and J is essentially an optimal control problem and that can be formulated by the following minimization problem: Find the solution $\phi \in U$ of

$$\begin{cases} D_t\phi & = & A(\phi) \\ \phi(t = t_0) & = & u^0 \\ \hat{u^0} & = & Arg \inf_{u^0} J(u^0, \phi), \end{cases} \tag{2.10}$$

where $J(u^0, \phi)$ is defined in (2.1) if we substitute $\phi(t)$ for $u(t)$ in (2.1). In the paper, we assume that the nonlinear evolution equation in (2.10) has a unique solution.

REMARK 2.6. *Since u^0 is the initial condition, $\phi = \phi(u^0)$. Hence, it is meaningful that $\hat{u^0} = Arg \inf_{u^0} J(u^0, \phi)$ in (2.10).*

By applying variational calculus techniques, we can show that the optimal control problem (2.10) is equivalent to the optimal system as form [4] stated by

$$\begin{cases} D_t\phi & = & A(\phi) \\ \phi(t = t_0) & = & u^0 \\ -D_t\phi^* & = & (\nabla A(\phi))^*\phi^* - \bar{H}^* R^{-1}(H\phi - \phi_{obs}) \\ \phi^*(t = T) & = & 0 \\ \phi^*(t = t_0) & = & B^{-1}(u^0 - u^B) \end{cases} \tag{2.11}$$

REMARK 2.7. *The fifth equation in (2.11) is the optimality condition.*

REMARK 2.8. *Let $S^A(t)$ denote an operator semigroup with generator A. By the notation of a semigroup,*

$$\phi(t) = S^A(t)u^0.$$

Hereafter we will use similar notations to denote semigroups. If A is a maximal dissipative operator on V, the exponential formula of the nonlinear semigroup [2] is

$$S^A(t)u^0 = \lim_{n\to\infty} (1 - \frac{t}{n}A)^{-n} u^0,$$

where the limit is taken in strong topology sense. Further,

$$\phi^*(t) = -\int_t^T S^{-(\nabla A(\phi))^*}(t-s)(\bar{H}^* R^{-1}(H\phi - \phi_{obs}))(s)ds, \tag{2.12}$$

where $S^{-(\nabla A(\phi))^*}t = e^{-(\nabla A(\phi))^* t}$ *since the generator* $-(\nabla A(\phi))^*$ *is a bounded linear operator.*

Next we briefly discuss the sensitivity analysis in a functional analysis framework. We still consider the nonlinear evolution equation

$$\left\{ \begin{array}{rcl} D_t\phi & = & A(\phi) \\ \phi(t=t_0) & = & u^0 \end{array} \right. \tag{2.13}$$

and a nonlinear functional $J(u^0, \phi)$. Define the tangent linear (perturbation) problem of the forward problem (2.13) by the form

$$\left\{ \begin{array}{rcl} D_t\delta\phi - (\nabla A(\phi))\delta\phi & = & 0 \\ \phi_0(t=t_0) & = & \delta u^0 \end{array} \right. \tag{2.14}$$

and the adjoint problem by the form

$$\left\{ \begin{array}{rcl} -D_t\phi^* & = & (\nabla A(\phi))^*\phi^* - \bar{H}^* R^{-1}(H\phi - \phi_{obs}) \\ \phi^*(t=T) & = & 0. \end{array} \right. \tag{2.15}$$

We want to know how sensitive the functional J is to the perturbation $\delta\phi$. Let $\delta_\phi J$ be the perturbation of $J(u^0, \phi)$ when $\delta\phi$ is nonzero. Then

$$\delta_\phi J = -(\delta u^0, \phi^*(t=t_0)). \tag{2.16}$$

The proof can be found in [4]. Equation (2.16) shows how the sensitivity of the functional $J(u^0, \phi)$ is related to initial condition u^0. From the above, we find that the solution ϕ^* of the adjoint problem (2.15) is responsible for the sensitivity of the functional J to the initial condition. It is often called an influence function or importance function [9].

Let ϕ^{**} be the solution of the second order adjoint problem by

$$\left\{ \begin{array}{rcl} -D_t\phi^{**} & = & (\nabla A(\phi))^*\phi^{**} + (\nabla^2 A(\phi)\delta\phi)^*\phi^* - \partial_\phi^2 J(\phi^0, \phi)\delta\phi \\ \phi^{**}(t=T) & = & 0, \end{array} \right. \tag{2.17}$$

where ϕ^* is the solution of the first order adjoint equation 2.15 and $\nabla^2 A(\phi)$ is the second derivative (Hessian) of A. Second order adjoint information is often used in data assimilation while applying numerical optimization algorithms [5]. Particularly, Hessian vector products are used in the computation of Hessian singular vectors in data assimilation [14].

3. Operator Splitting Methods. In this subsection we show that *Hypothesis 1* can be realized using a constructive mannerism and error estimates are given.

If $A(\phi)$ in (2.13) has a complicated structure and consists of different parts, e.g., $A(\phi)$ consists of an advection operator, a diffusion operator and a reaction operator, splitting methods are advocated for solving (2.13) [6]. The basic idea behind operator splitting is to reduce a complicated problem into smaller or simpler subproblems such that different parts can be solved efficiently with appropriate integration formulas.

Here we will use the Marchuk-Strang symmetrical multi-component splitting [8, 13] to solve (2.13).

Let $A(\phi) = \sum_{j=1}^{J} A_j(\phi)$, where $\mathcal{A}_j$ is the Lie operator associated with each operator A_j ($\mathcal{A}$ is associated with A). These Lie operators $\mathcal{A}_j$ (or $\mathcal{A}$) are linear operators on the space of operators acting on the solution space U of (2.13). For any $v \in U$ and any operator g on U, by definition of a Lie operator [6], it follows that

$$\mathcal{A}g(v) = g'(v)A(v).$$

So for the solution $\phi(t)$ of (2.13),

$$\mathcal{A}g(\phi(t)) = g'(\phi(t))A(\phi(t)) = \frac{\partial}{\partial t} g(\phi(t)).$$

Let I be identity operator. Using Lie-Taylor series [3] gives us

$$\phi(t+\tau) = (e^{\tau \mathcal{A}} I)\phi(t).$$

We split the problem (2.13) into J subproblems, $D_t\phi_j = A_j(\phi_j)$, $j = 1, \cdots, J$. We apply the Marchuk-Strang symmetrical multi-component splitting over time intervals $[t_k, t_{k+1}]$, where $t_{k+1} = t_k + \tau$ with constant time step length τ, to obtain

$$\left\{ \begin{array}{rcl} D_t\phi_1 & = & A_1(\phi_1), \quad \phi_1(t_k) = \phi_1'(t_k), \quad t \in [t_k, t_k + \frac{\tau}{2}] \\ & \cdots & \\ D_t\phi_{J-1} & = & A_{J-1}(\phi_{J-1}), \quad \phi_{J-1}(t_k) = \phi_{J-2}(t_k + \frac{\tau}{2}), \quad t \in [t_k, t_k + \frac{\tau}{2}] \\ D_t\phi_J & = & A_J(\phi_J), \quad \phi_J(t_k) = \phi_{J-1}(t_k + \frac{\tau}{2}), \quad t \in [t_k, t_k + \tau] \\ D_t\phi_{J-1}' & = & A_{J-1}(\phi_{J-1}'), \quad \phi_{J-1}'(t_k + \frac{\tau}{2}) = \phi_J(t_{k+1}), \quad t \in [t_k + \frac{\tau}{2}, t_{k+1}] \\ & \cdots & \\ D_t\phi_1' & = & A_1(\phi_1'), \quad \phi_1(t_k + \frac{\tau}{2}) = \phi_2'(t_{k+1}), \quad t \in [t_k + \frac{\tau}{2}, t_{k+1}]. \end{array} \right. \tag{3.1}$$

Let $\mathcal{S}_{j,\frac{1}{2}\tau} = e^{\frac{1}{2}\tau \mathcal{A}_j}$, $\mathcal{S}_{J,\tau} = e^{\tau \mathcal{A}_J}$, and ϕ_k be approximations of $\phi(t_k)$, $k = 1, \cdots, N$. Applying the Baker-Campbell-Hausdorf formula of a Lie operator [6] gives us

$$\phi(t_{k+1}) = \mathcal{S}_{1,\frac{1}{2}\tau} \cdots \mathcal{S}_{J-1,\frac{1}{2}\tau} \mathcal{S}_{J,\tau} \mathcal{S}_{J-1,\frac{1}{2}\tau} \cdots \mathcal{S}_{1,\frac{1}{2}\tau} I \phi(t_k) + O(\tau^3) \tag{3.2}$$

and

$$\phi_{k+1} = \mathcal{S}_{1,\frac{1}{2}\tau} \cdots \mathcal{S}_{J-1,\frac{1}{2}\tau} \mathcal{S}_{J,\tau} \mathcal{S}_{J-1,\frac{1}{2}\tau} \cdots \mathcal{S}_{1,\frac{1}{2}\tau} I \phi_k. \tag{3.3}$$

Let $\mathcal{S}_{[t_k,t_{k+1}]} = \mathcal{S}_{1,\frac{1}{2}\tau} \cdots \mathcal{S}_{J-1,\frac{1}{2}\tau} \mathcal{S}_{J,\tau} \mathcal{S}_{J-1,\frac{1}{2}\tau} \cdots \mathcal{S}_{1,\frac{1}{2}\tau}$ be the operator splitting procedure over $[t_k, t_{k+1}]$. Then

$$\phi_N = \Pi_{k=0}^{N-1} \mathcal{S}_{[t_k,t_{k+1}]} I \phi_0. \tag{3.4}$$

REMARK 3.1. *By means of the Lie operator formalism, we transform a nonlinear splitting into the compositions of linear operators and so Hypothesis 1 is realized.*

REMARK 3.2. *The term $O(\tau^3)$ in (3.2) represents the leading term of the local splitting error. It is a second order splitting scheme in the time because $\tau^{-1}\|\phi(t_{k+1}) - \phi(t_k)\| = O(\tau^2)$. The symmetrical operator splitting scheme has second order consistency. When $A_j(\phi)$ and $A_l(\phi)$ commute each other, i.e., for any $j \neq l$, $A_j' A_l = A_j A_l'$, where A_j' is the derivative with regard to ϕ, then no splitting error occurs [6].*

REMARK 3.3. *In fact, (3.4) gives rise to a general Strang's product formula*

$$\lim_{n\to\infty} [e^{\frac{t}{2n}\mathcal{A}_1} \cdots e^{\frac{t}{2n}\mathcal{A}_{J-1}} e^{\frac{t}{n}\mathcal{A}_J} e^{\frac{t}{2n}\mathcal{A}_{J-1}} \cdots e^{\frac{t}{2n}\mathcal{A}_1}]^n I u^0 = e^{t\mathcal{A}} I u^0.$$

REMARK 3.4. *If all A_j's are linear on ϕ, then we can represent the numerical solution of (2.13) by linear operator semi-groups,*

$$\phi_{k+1} = S^{A_1}(\frac{1}{2}\tau)\cdots S^{A_{J-1}}(\frac{1}{2}\tau)S^{A_J}(\tau)S^{A_{J-1}}(\frac{1}{2}\tau)\cdots S^{A_1}(\frac{1}{2}\tau)\phi_k,$$

where $S^{A_i}(\frac{1}{2}\tau)$ $(i = 1, \cdots, J)$ denotes the operator semigroup and A_i are the corresponding generators. Here it only requires that A_i, $i = 1, \cdots, J$, are closed, densely defined linear operators, but can be unbounded. By the exponential formula of linear semigroups [2],

$$S^{A_i}(\frac{1}{2}\tau) = \lim_{n\to\infty} e^{A_i(I-\frac{A_i}{n})^{-1}\frac{1}{2}\tau}, \quad i = 1, \cdots J-1,$$

and the limit is taken in strong topology sense. Similarly we can find the exponential formula for $S^{A_J}(t)$.

We can also use operator splitting to solve the perturbation equation (2.14) such that

$$\delta\phi_{k+1} = \bar{\mathcal{S}}_{1,\frac{1}{2}\tau}\cdots\bar{\mathcal{S}}_{J-1,\frac{1}{2}\tau}\bar{\mathcal{S}}_{J,\tau}\bar{\mathcal{S}}_{J-1,\frac{1}{2}\tau}\cdots\bar{\mathcal{S}}_{1,\frac{1}{2}\tau}\delta\phi_k := \bar{\mathcal{S}}_{[t_k,t_{k+1}]}\delta\phi_k$$

and

$$\delta\phi_N = \Pi_{k=0}^{N-1}\bar{\mathcal{S}}_{[t_k,t_{k+1}]}\delta\phi_0,$$

where $\bar{\mathcal{S}}_{j,\frac{1}{2}\tau} = S^{\nabla A_j(\phi)}(\frac{1}{2}\tau) = e^{\frac{1}{2}\tau\nabla A_j(\phi)}$, $j = 1, \cdots, J-1$, and $\bar{\mathcal{S}}_{J,\tau} = S^{\nabla A_J(\phi)}(\tau) = e^{\tau\nabla A_J(\phi)}$ since $\nabla A_j(\phi)$'s are bounded linear operators.

We split the problem (2.15) into J subproblems, $-D_t\phi_j^* = (\nabla A_j(\phi))^*\phi_j^*$, $j = 1, \cdots, J$. We apply the Marchuk-Strang symmetrical multi-component splitting over a time interval $[t_{k+1}, t_k]$ to get

$$\left\{\begin{array}{rcl} -D_t\phi_1^* &=& (\nabla A_1(\phi))^*\phi_1^*, \quad \phi_1^*(t_{k+1}) = \phi_1^{*'}(t_{k+1}), \quad t\in[t_{k+1}, t_k+\frac{\tau}{2}] \\ & \cdots & \\ -D_t\phi_{J-1}^* &=& (\nabla A_{J-1}(\phi))^*\phi_{J-1}^*, \quad \phi_{J-1}^*(t_{k+1}) = \phi_{J-2}^*(t_{k+1}-\frac{\tau}{2}), \quad t\in[t_{k+1}, t_k+\frac{\tau}{2}] \\ -D_t\phi_J^* &=& (\nabla A_J(\phi))^*\phi_J^*, \quad \phi_J(t_{k+1})^* = \phi_{J-1}^*(t_{k+1}-\frac{\tau}{2}), \quad t\in[t_{k+1}, t_k] \\ -D_t\phi_{J-1}^{*'} &=& (\nabla A_{J-1}(\phi))\phi_{J-1}^{*'}, \quad \phi_{J-1}^{*'}(t_k+\frac{\tau}{2}) = \phi_J^*(t_k), \quad t\in[t_k+\frac{\tau}{2}, t_k] \\ & \cdots & \\ -D_t\phi_1^{*'} &=& (\nabla A_1(\phi))^*\phi^{*'}, \quad \phi_1^{*'}(t_k+\frac{\tau}{2}) = \phi_2^{*'}(t_k), \quad t\in[t_k+\frac{\tau}{2}, t_k]. \end{array}\right. \tag{3.5}$$

Thus, we obtain

$$\bar{\mathcal{S}}^*_{j,\frac{1}{2}\tau} = e^{-\frac{1}{2}\tau(\nabla A_j(\phi))^*}, \quad (j = 1, \cdots, J-1), \quad \bar{\mathcal{S}}^*_{J,\tau} = e^{-\tau(\nabla A_J(\phi))^*} \tag{3.6}$$

and

$$\left\{\begin{array}{rcl} \phi_k^* &=& \bar{\mathcal{S}}^*_{[t_{k+1},t_k]}\phi_{k+1}^* + \partial_{\phi_k}J(u^0,\phi) \\ &=& \bar{\mathcal{S}}^*_{1,-\frac{1}{2}\tau}\cdots\bar{\mathcal{S}}^*_{J-1,-\frac{1}{2}\tau}\bar{\mathcal{S}}^*_{J,-\tau}\bar{\mathcal{S}}^*_{J-1,-\frac{1}{2}\tau}\cdots\bar{\mathcal{S}}^*_{1,-\frac{1}{2}\tau}\phi_{k+1}^* + \partial_{\phi_k}J(u^0,\phi) \\ \phi_N^* &=& 0, \end{array}\right. \tag{3.7}$$

where $\partial_{\phi_k}J$ be the partial differential of $J(u^0,\phi)$ with respect to ϕ_k for a fixed first argument u^0 and $\phi = \{\phi_0, \cdots, \phi_N\}$.

Similarly, we use the Marchuk-Strang symmetric splitting for the second order adjoint problem (2.17), we can obtain that [15]

$$\left\{\begin{array}{lcl}\phi_k^{**} & = & \bar{\mathcal{S}}^*_{[t_{k+1},t_k]}\phi^{**}_{k+1} + (\bar{\bar{\mathcal{S}}}_{[t_k,t_{k+1}]}\delta\phi_k)^*\phi^*_{k+1} + \partial^2_{\phi_k}J(u^0,\phi)\delta\phi_k \\ \phi^{**}_N & = & 0,\end{array}\right. \tag{3.8}$$

where $\bar{\bar{\mathcal{S}}}_{[t_k,t_{k+1}]} = \bar{\bar{\mathcal{S}}}_{1,\frac{1}{2}\tau}\cdots\bar{\bar{\mathcal{S}}}_{J-1,\frac{1}{2}\tau}\bar{\bar{\mathcal{S}}}_{J,\tau}\bar{\bar{\mathcal{S}}}_{J-1,\frac{1}{2}\tau}\cdots\bar{\bar{\mathcal{S}}}_{1,\frac{1}{2}\tau}$ and $\bar{\bar{\mathcal{S}}}_{j,\frac{1}{2}\tau} = e^{\frac{1}{2}\tau\nabla^2 A_j(\phi)}$, $j = 1,\cdots,J-1$, $\bar{\bar{\mathcal{S}}}_{J,\tau} = e^{\tau\nabla^2 A_J(\phi)}$.

Let $\nabla^\tau_{u^0}J(u^0,\phi(u^0))$ be the approximation of $\nabla_{u^0}J(u^0,\phi(u^0))$ by the operator splitting method and $\nabla^{2,\tau}_{u^0}J(u^0,\phi(u^0))$ the approximation of $\nabla^2_{u^0}J(u^0,\phi(u^0))$. Then we have the following theorem.

THEOREM 3.1. *Let ϕ^* be defined in (3.7) and ϕ^{**} be defined in (3.8). Then Marchuk-Strang operator splitting gives rise to*

$$\begin{aligned}\nabla^\tau_{u^0}J(u^0,\phi(u^0)) &= \phi^*_0 + B^{-1}(u^0-u^B)\\ \nabla^{2,\tau}_{u^0}J(u^0,\phi(u^0))\delta u^0 &= \phi^{**}_0 + B^{-1}\delta u^0,\end{aligned} \tag{3.9}$$

where ϕ^*_0 is the operator splitting solution at t_0 defined in (3.7) and ϕ^{**}_0 is the operator splitting solution at t_0 defined in (3.8).

Proof. Since $\partial_{\phi_k}J = \bar{H}^*_k R_k^{-1}(H_k\phi_k - \phi_k^{obs})$, it follows from (2.8) that

$$\begin{aligned}\nabla^\tau_{u^0}J &= B^{-1}(u^0-u^B) + \sum_{k=0}^N \bar{\mathcal{S}}^*_{[t_k,t_0]}\partial_{\phi_k}J\\ &= B^{-1}(u^0-u^B) + \sum_{k=0}^N \bar{\mathcal{S}}^*_{[t_1,t_0]}\cdots\bar{\mathcal{S}}^*_{[t_k,t_{k-1}]}\partial_{\phi_k}J\\ &= B^{-1}(u^0-u^B) + I\partial_{u^0}J\\ &+ \bar{\mathcal{S}}^*_{[t_1,t_0]}(\partial_{\phi_1}J + \bar{\mathcal{S}}^*_{[t_2,t_1]}(\partial_{\phi_2}J + \cdots + \bar{\mathcal{S}}^*_{[t_{N-1},t_{N-2}]}(\partial_{\phi_{N-1}}J + \bar{\mathcal{S}}^*_{[t_N,t_{N-1}]}\partial_{\phi_N}J))).\end{aligned} \tag{3.10}$$

By the recurrence definition of ϕ^*_k in (3.7), it follows that

$$\nabla^\tau_{u^0}J = \phi^*_0 + B^{-1}(u^0-u^B).$$

This verifies the first equation in (3.9). The proof of the second equation in (3.9) is similar to the first equation. □

REMARK 3.5. *By Theorem 3.1, we use the solutions of the first order adjoint equation (3.7) to evaluate the gradient of the cost functional and the second order adjoint (3.8) to compute Hessian vector products.*

From Theorem 3.1 and the operator splitting procedure, we compute $J(u^0)$ by solving an adjoint problem that depends on the forward trajectory.

As for stability of the operator splitting scheme, if we have

$$\|e^{\frac{1}{2}\tau\mathcal{A}_j}\| \le e^{\frac{1}{2}\tau\omega_j},\quad j=1,\cdots,j-1,\quad \text{and}\quad \|e^{\tau\mathcal{A}_J}\| \le e^{\tau\omega_J}, \tag{3.11}$$

then $\|\phi_{k+1}\| \le e^{\tau\omega}\|\phi_k\|$, where $\omega = \sum_{j=1}^J \omega_j$. Hence, the stability holds on any finite time interval $[t_0,T]$ if $\omega > 0$. It also holds for an arbitrary large time interval if $\omega \le 0$. In practice the operator splitting scheme is stable if each sub-step is stable. By Lax's equivalence theorem [7], consistency and stability together imply convergence, and higher order consistency yields faster convergence. In particular, we have the following theorem for the global splitting error:

THEOREM 3.2. *Let* $\mu(\nabla A(\phi)) := \lim_{h\to 0+} \frac{\|I+h\nabla A(\phi)\|-1}{h} \le \lambda$, *then*

$$\|\phi(t_n) - \phi_n\| \le C\tau^3(e^{n\lambda\tau} - 1)(e^{\lambda\tau} - 1)^{-1},$$

where C *is a positive number independent of* τ.

Proof. If the perturbation of the initial condition of (2.13) is δu^0, then by the perturbation equation (2.14) and using a semigroup expression, we have

$$\delta\phi(t) = e^{t\nabla A(\phi)}\delta u^0.$$

Consequently,

$$\|\delta\phi(t)\| \le \|e^{t\nabla A(\phi)}\|\|\delta u^0\| \le e^{\mu(\nabla A(\phi))}\|\delta u^0\| \le e^{\lambda t}\|\delta u^0\|, \tag{3.12}$$

where we have used Proposition 2.1 in [12] in the second step. By (3.2), the local splitting errors do not exceed $C\tau^3$ for some constant C. In computing ϕ_2 there is an error of $C\tau^3$ in the initial condition, and by (3.12), the effect of this error at t_2 is $C\tau^3 e^{\lambda\tau}$. Thus, the global splitting error at t_2 is $C\tau^3 + C\tau^3 e^{\lambda\tau}$. Similarly the global splitting error at t_3 is

$$C\tau^3 + (C\tau^3 + C\tau^3 e^{\lambda\tau})e^{\lambda\tau}.$$

Repeating the procedure in the same way we find that the global splitting error at t_n is

$$\sum_{k=1}^{n} C\tau^3 e^{(n-k)\lambda\tau} = C\tau^3 \sum_{k=0}^{n-1} e^{k\lambda\tau} = C\tau^3(e^{n\lambda\tau} - 1)(e^{\lambda\tau} - 1)^{-1}.$$

□

REMARK 3.6. *The* $\mu(\nabla A(\phi))$ *defined in Theorem 3.2 is called a logarithmic norm [12] of the bounded linear operator* $\nabla A(\phi)$.

REMARK 3.7. *Since* $e^z - 1 = O(z)$, *we find that* $(e^{n\lambda\tau} - 1)(e^{\lambda\tau} - 1)^{-1} = O(\frac{1}{\tau})$, *and so* $\|\phi(t_n) - \phi_n\| = O(\tau^2)$. *Hence, the global Marchuk-Strang splitting error is second order for time step of length* τ.

By utilizing Theorem 3.2 and Remark 3.7, the following proposition follows immediately.

PROPOSITION 3.3. *Let* $\nabla^\tau_{u^0} J(u^0, \phi(u^0))$ *be the approximation of* $\nabla_{u^0} J(u^0, \phi(u^0))$ *by the Marchuk-Strang operator splitting method. If* $\mu(\nabla A(\phi))$ *and* $\mu((\nabla A(\phi))^*)$ *are bounded, then*

$$\|\nabla_{u^0} J(u^0, \phi(u^0)) - \nabla^\tau_{u^0} J(u^0, \phi(u^0))\| \le C\tau^2,$$

where C *is a positive number independent of* τ.

Similarly we can estimate the local and global splitting error for $\|\nabla^2_{u^0} J(u^0, \phi(u^0))v - \nabla^{2,\tau}_{u^0} J(u^0, \phi(u^0))v\|$ for any user-defined element v.

4. Conclusions. In this paper, we presented a framework for 4D variational data assimilation and discuss the Marchuk-Strang symmetrical operator splitting methods in the functional 4D-Var. The proposed analysis stems from the rapid theoretical advance of 4D-Var and the desire to translate it into real life applications.

First, we investigated a general framework of 4D-Var in the setting of functional operators. Constructing and solving a adjoint problem gives rise to an efficient approach to evaluate the gradient of the cost functional with respect to an unknown parameter. This framework help us understand 4D-Var theoretically.

Second, we explored symmetrical operator splitting methods, which are used to implement a 4D-Var problem in practice. It concluded that a large and complicated 4D-Var problem can be split into many subproblems and the Marchuk-Strang symmetrical multi-component splitting method gives rise to a second order splitting error for time step.

REFERENCES

[1] V. Akcelik, G. Biros, A. Draganescu, O. Ghattas, J. Hill, and B. v. Bloemen-Waanders, *Inversion of Airborne Contaminants in a Regional Model*, Lecture Notes in Computer Science series, vol. 3993, pp. 481-488, Springer-Verlag, Heidelberg, 2006.

[2] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Editura Academiei, Bucuresti, 1976.

[3] W. Hundsdorfer and J. G. Verwer, *Numerical Solutions of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer-Verlag, New York, 2003.

[4] L. Jiang and C. C. Douglas, *A Unified Aanalysis of 4D Variational Data Assimilation and Its Application*, submitted to Computing.

[5] F.X. Le Dimet, I.M. Navon, and D. Daescu, *Second Order Information in Data Assimilation*, Monthly Weather Review, 130(3) (2002), pp. 629-648.

[6] D. Lanser and J. G. Verwer, *Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling*, J. Comput. Appl. Math. 111 (1999), pp. 201–216.

[7] P. Lax, *Functional Analysis*, Wiley Interscience, New York, 2002.

[8] G. Marchuk, *Adjoint Equations and Analysis of Complex System*, Mathematics and Its Applications Series 29, Springer-Verlag, New York, 1995.

[9] G. Marchuk, V. Shutyaev and G. Bocharov, *Adjoint Equations and Analysis of Complex System: Application to Virus Infection Modelling*, J. Computational and Applied Math, 184 (2005), pp. 177-204.

[10] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997.

[11] F. A. Rihan and C. G. Collier, *Four-Dimensional Data Assimilation and Numerical Weather Prediction*, CERS Data Assimilation Technical Report No. 1 (2003), University of Salford, UK.

[12] G. Söderlind, *The logarithmic norm. History and modern theory*, Bit., Vol. 46 (no.3) (2006), pp. 631-652.

[13] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal. 5 (1968), pp. 506–517.

[14] Z. Wang, K. Droegemeier, and L. White, *The Adjoint Newton Algorithm for Large-Scale Unconstrained Optimization in Meteorology Applications*, Comput. Optim. Appl., 10(3) (1998), pp. 283-320.

[15] L. Zhang, Large-Scale Simulations Using First and Second Order Adjoints with Applications in Data Assimilation, Virginia Polytechnic Institute and State University, thesis, 2007.

[16] L. Zhang and A. Sandu, *Data Assimilation in Multiscale Chemical Transport Models*, Lecture Notes in Computer Science series, vol. 4487, pp. 1026-1033, Springer-Verlag, Heidelberg, 2007.