Carnegie Mellon University

A short tutorial on the Adjoint Method

Varun Shankar

Department of Mechanical Engineering

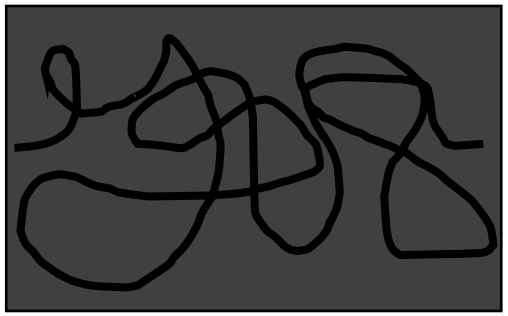
Motivation

- Computing gradients

$$L = MSE(u, \hat{u}(p))$$

$$\frac{dL}{dp} ?$$

In → $f \circ g \circ h \circ i \circ j$ → Out



In → → Out

$$u = f_p(g_p(h_p(\dots x)) \dots$$

$$g(u,p) = 0 \quad = p_1 \frac{du}{dx} + p_1 \frac{d^2u}{dx^2} - f(p_3, x)$$

Derivation

$$g(u,p) = 0 \qquad\qquad L(u,p)$$

$$\frac{dL}{dp} = \frac{\partial L}{\partial u}\frac{\partial u}{\partial p} + \frac{\partial L}{\partial p} = L_u u_p + L_p$$

$$dg = g_u u_p + g_p = 0$$

$$u_p = -g_u^{-1} g_p$$

Tangent Linear

NxN Nx1

Derivation

$$g(u,p) = 0 \qquad\qquad L(u,p)$$

$p \rightarrow g \rightarrow u \rightarrow L$

$$\frac{dL}{dp} = \underset{\text{LxN}}{-L_u} \overset{\text{NxP}}{(g_u^{-1} g_p)} + L_p \qquad\qquad u_p = -g_u^{-1} g_p$$

$$\frac{dL}{dp} = \underbrace{(-L_u g_u^{-1})}_{\lambda^T} g_p + L_p$$

A Linear Example

$$g(u,p) = \boldsymbol{K}\boldsymbol{u} - \boldsymbol{f} = 0 \qquad L(u,p) = \mathrm{MSE}(u)$$

$$\frac{dL}{d\boldsymbol{p}} = \boldsymbol{\lambda}^T \boldsymbol{g_p} + L_{\boldsymbol{p}}$$

$$\boldsymbol{\lambda}^{\mathrm{T}} = -L_{\boldsymbol{u}} \boldsymbol{g_u}^{-1}$$

$$\boldsymbol{g_u^T \lambda} = -L_{\boldsymbol{u}}^T$$

$$\frac{dL}{d\boldsymbol{p}} = \boldsymbol{\lambda}^T \boldsymbol{g_p}$$

$$\boldsymbol{K}^T \boldsymbol{\lambda} = -L_{\boldsymbol{u}}^T$$

$$\frac{dL}{d\boldsymbol{p}} = \boldsymbol{\lambda}^T (\boldsymbol{K_p u} - \boldsymbol{f_p})$$

Continuous Adjoint

$$\boldsymbol{K u = f}$$

$$\boldsymbol{K^T \lambda = b} \qquad \boldsymbol{\lambda^T K = b^T}$$

$$\boldsymbol{\lambda^T K u = b^T u}$$

$$\boxed{\boldsymbol{\lambda^T f = b^T u}}$$

Continuous Adjoint

$$Lu = f$$

$$L^*\lambda = b$$

$$(\lambda, u) = \int_\Omega \lambda^T u$$

$$(\lambda, f) = b^T u$$

$$(\lambda, Lu) = (L^*\lambda, u)$$

Continuous Adjoint Example

$$f = Lu = u' - u''$$

$$\frac{\partial L}{\partial u} = L^*\lambda = ?\ -\lambda' - \lambda''$$

Use the same forward solver

$$(\lambda, Lu) = \int \lambda(u' - u'')$$

$$\vdots$$

$$= \int (\underbrace{-\lambda' - \lambda''}_{L^*\lambda})u \quad + BCs$$

$$(\lambda, Lu) = (L^*\lambda, u)$$

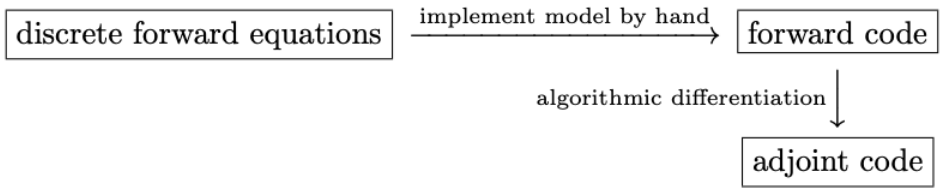


Fig. 1.1: The traditional approach to developing adjoint models. The forward model is implemented by hand, and its adjoint derived either by hand or with the assistance of an algorithmic differentiation tool.

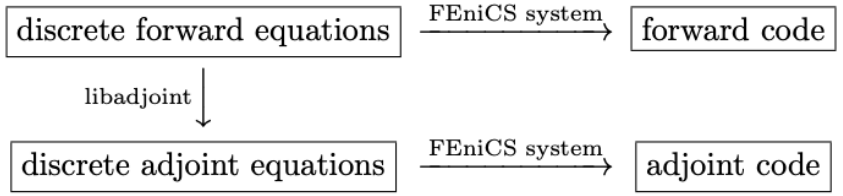


Fig. 1.2: The approach to adjoint model development advocated in this paper. The user specifies the discrete forward equations in a high-level language similar to mathematical notation; the discrete forward equations are explicitly represented in memory in the UFL format [2, 1]. libadjoint automatically derives the corresponding in-memory representation of the discrete adjoint equations, from the in-memory representation of the forward problem. Both the forward and adjoint equations are then passed to the FEniCS system, which automatically generates and executes the code necessary to compute the forward and adjoint solutions.

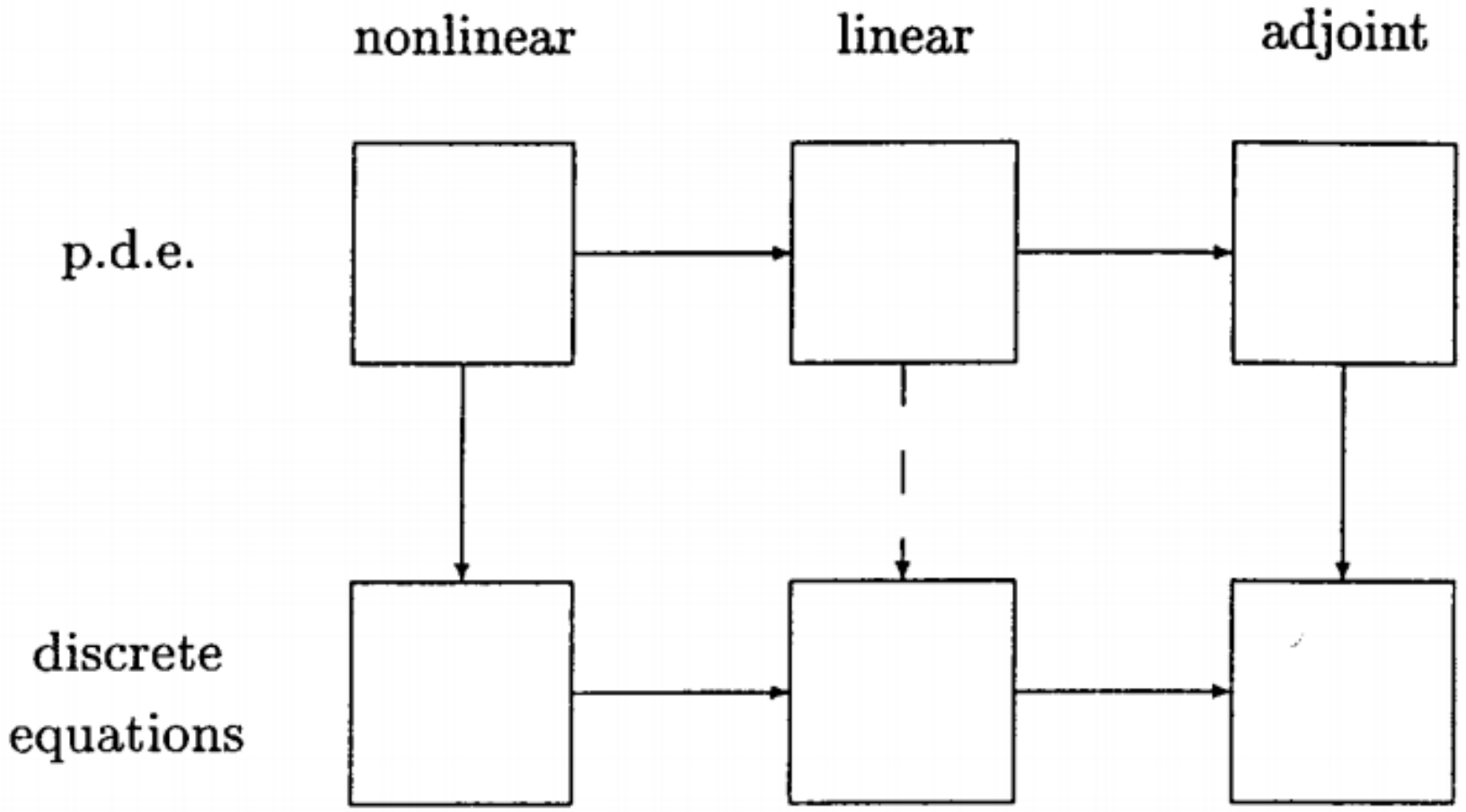


Figure 1. Alternative approaches to forming discrete adjoint equations.

http://www.piercelab.caltech.edu/assets/papers/ftc00.pdf

Time Dependent

$$g(u,p) = 0$$

$$g(u,p) = \dot{u}$$

$$L(u,p)$$

$$L(u,p) = \int_0^T l(u,p,t)dt$$

$$g_u^T \lambda = -L_u^T$$

$$l_u + \lambda^T g_u = \dot{\lambda}^{\mathrm{T}}$$

$$\lambda(T) = 0$$

$$\frac{dL}{dp} = \lambda^{\mathrm{T}} g_p + L_p$$

$$\frac{dL}{dp} = \int_0^T \lambda^{\mathrm{T}} g_p + l_p dt + \lambda^{\mathrm{T}} u_p \Big|_{t=0}$$