Adjoint-based optimization of external force field in OPAL

Mohsen Sadr[1], Andreas Adelmann[1], Dan T. Abell[2], Thomas M. Antonsen[3], and Irving Haber[3]

[1]Paul Scherrer Institute, Forschungsstrasse 111, CH-5232 Villigen, Switzerland
[2]RadiaSoft LLC, Boulder, Colorado, USA
[3]Department of Physics, University of Maryland at College Park, USA

Parameter Optimization for a Particle System

Consider the evolution of single particle distribution function

$$\mathcal{L}[f] := \frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\nabla_{\boldsymbol{x}}(f) + \nabla_{\boldsymbol{x}}\phi\cdot\nabla_{\boldsymbol{v}} f = 0\ , \tag{1}$$

where $\phi(\boldsymbol{x},t) = \phi_{\text{self}}(\boldsymbol{x},t) + \phi_{\text{ext.}}(\boldsymbol{x};\boldsymbol{\alpha})$ and

$$\nabla^2\phi_{\text{self}} = -\int q f d\boldsymbol{v}\ . \tag{2}$$

The goal is to find parameters $\boldsymbol{\alpha}$ of the external potential $\phi_{\text{ext.}}(\boldsymbol{x};\boldsymbol{\alpha})$ such that given an initial distribution $f_0 = f(\boldsymbol{v};\boldsymbol{x},t=t_0)$, the simulation leads to the final time distribution $f_\tau = f(\boldsymbol{v};\boldsymbol{x},t=\tau)$.

Adjoint equation in Hamiltonian Dynamics

Consider the Hamiltonian

$$H(\boldsymbol{x},\boldsymbol{p},t) = mc^2\gamma + q\phi, \tag{3}$$

where c is the speed of light, ϕ the scalar potential, γ the relativistic Lorentz factor

$$\gamma = \sqrt{1 + \tfrac{1}{(mc)^2}(\boldsymbol{p} - q\boldsymbol{A})^2}, \tag{4}$$

with the equations of motion

$$\frac{d\boldsymbol{x}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}} \quad \text{and} \quad \frac{d\boldsymbol{p}}{dt} = \frac{\partial H}{\partial \boldsymbol{x}}\ . \tag{5}$$

Adjoint Relation:

The idea is to find a relation between perturbation of a loss function with respect to design parameter $\boldsymbol{\alpha}$ denoted by $(.)^{(X)}$ and perturbation of the coordinates at final time $(.)^{(Y)}$ given a figure of merit C.

First, let us perturb the equations of motion, i.e.

$$\frac{d\delta\boldsymbol{x}_j}{dt} = \delta\boldsymbol{x}_j\cdot\frac{\partial^2 H}{\partial\boldsymbol{x}\partial\boldsymbol{p}} + \delta\boldsymbol{p}_j\cdot\frac{\partial^2 H}{\partial\boldsymbol{p}^2} - q_j\frac{\partial}{\partial\boldsymbol{p}}(\boldsymbol{v}_j\cdot\delta\boldsymbol{A}), \tag{6}$$

$$\frac{d\delta\boldsymbol{p}_j}{dt} = -\delta\boldsymbol{x}_j\cdot\frac{\partial^2 H}{\partial\boldsymbol{x}^2} - \delta\boldsymbol{p}_j\cdot\frac{\partial^2 H}{\partial\boldsymbol{p}\partial\boldsymbol{x}} - q_j\frac{\partial\delta\phi}{\partial\boldsymbol{x}} + q_j\frac{\partial}{\partial\boldsymbol{x}}(\boldsymbol{v}_j\cdot\delta\boldsymbol{A}). \tag{7}$$

Next, construct a symplectic area integrated over particles and till time τ via

$$\sum_j\left(\delta\boldsymbol{x}_j^{(Y)}\cdot\delta\boldsymbol{p}_j^{(X)} - \delta\boldsymbol{x}_j^{(X)}\cdot\delta\boldsymbol{p}_j^{(Y)}\right)\Big|_0^\tau = \int_0^\tau dt\int d\boldsymbol{x}\left\{\left[\delta\rho^{(X)}\delta\phi^{(Y)} - \delta\boldsymbol{J}^{(X)}\cdot\delta\boldsymbol{A}^{(Y)}\right] - (Y\leftrightarrow X)\right\}\ . \tag{8}$$

In the electrostatic setting with determined boundary conditions, we find

$$\sum_j\left(\delta\boldsymbol{x}_j^{(Y)}\cdot\delta\boldsymbol{p}_j^{(X)} - \delta\boldsymbol{x}_j^{(X)}\cdot\delta\boldsymbol{p}_j^{(Y)}\right)\Big|_0^\tau = -\sum_k\delta\alpha_k\int_0^\tau dt\int d\boldsymbol{x}\left\{\delta\rho^{(A)}\frac{\partial\phi_{\text{ext.}}}{\partial\alpha_k}\right\}. \tag{9}$$

Since the total variation of the loss function has the form

$$\delta C = \sum_j \delta\boldsymbol{x}_j^{(X)}\frac{\partial F}{\partial\boldsymbol{x}_j} + \delta\boldsymbol{p}_j^{(X)}\frac{\partial F}{\partial\boldsymbol{p}_j}\ , \tag{10}$$

we note that the LHS of eq. (9) turns into δC by setting

$$\delta\boldsymbol{p}_j^{(Y)} = -\frac{\partial C}{\partial\boldsymbol{x}_j} \quad \text{and} \quad \delta\boldsymbol{x}_j^{(Y)} = \frac{\partial C}{\partial\boldsymbol{p}_j}\ . \tag{11}$$

Given $\phi_{\text{ext.}} = \phi_{\text{ext.}}(\boldsymbol{x};\boldsymbol{\alpha})$, the variation of the loss function (figure of merit) in terms of design parameters $\boldsymbol{\alpha}$ becomes

$$\delta C = -\sum_k\delta\alpha_k\int_0^\tau dt\int d\boldsymbol{x}\left\{\delta\rho^{(Y)}\frac{\partial\phi_{\text{ext.}}}{\partial\alpha_k}\right\}\ . \tag{12}$$

The gradient of the loss functions with respect to parameters is

$$\delta C/\delta\alpha_k = -\int_0^\tau dt\int d\boldsymbol{x}\left\{\rho^{(Y)}\frac{\partial\phi_{\text{ext.}}}{\partial\alpha_k}\right\} + \int_0^\tau dt\int d\boldsymbol{x}\left\{\rho\frac{\partial\phi_{\text{ext.}}}{\partial\alpha_k}\right\}\ , \tag{13}$$

which needs to be computed in the backward simulation.

Adjoint Relation from Distribution Perspective

Here, we find the adjoint equation using the kinetic equation. The idea is to formulate the problem as the unconstrained optimization problem with the loss functional

$$\mathcal{C} = \sum_l\frac{1}{2}D_l^2 + \int\int\int\gamma(\boldsymbol{v},\boldsymbol{x},t)\mathcal{L}[f(\boldsymbol{v},\boldsymbol{x},t)]d\boldsymbol{v}d\boldsymbol{x}dt \tag{14}$$

$$\text{where}\quad \boldsymbol{D} = \int\int\boldsymbol{H}(\boldsymbol{v},\boldsymbol{x})f(\boldsymbol{v},\boldsymbol{x},\tau)d\boldsymbol{v}d\boldsymbol{x} - \boldsymbol{\mu}_\tau \tag{15}$$

$$\text{and}\quad \boldsymbol{\mu}_\tau = \int\int\boldsymbol{H}(\boldsymbol{v},\boldsymbol{x})f_{\text{target}}(\boldsymbol{v},\boldsymbol{x})d\boldsymbol{v}d\boldsymbol{x}\ . \tag{16}$$

By letting the variational derivatives of $\mathcal{C}$ with respect to f to zero, we reach a backward equation for the Lagrange multipliers known as adjoint equation

$$\frac{\delta\mathcal{C}}{\delta f} = 0 \implies \mathcal{A}[\gamma] := \partial_t\gamma + \boldsymbol{v}\cdot\nabla_{\boldsymbol{x}}\gamma + \nabla_{\boldsymbol{v}}\gamma\cdot\nabla_{\boldsymbol{x}}\Big(\phi + \phi_{\text{ext.}}\Big) = 0 \tag{17}$$

with the final condition

$$\frac{\delta\mathcal{C}}{\delta f(\boldsymbol{v},\boldsymbol{x},\tau)} = 0 \implies \gamma(\boldsymbol{v},\boldsymbol{x},\tau) = \sum_l H_l D_l\ . \tag{18}$$

We note that the adjoint equation resembles the forward equation with the difference that it has a final time condition. Hence, it follows the same dynamics as the forward model. Moreover, the adjoint equation is conservative with respect to γ on particle trajectories.

By setting the perturbation of coordinates at final time as the Lagrange multipliers, then objective function C on the particle trajectory becomes

$$C = \sum_{i=1}^{\infty}\sum_l\frac{1}{2}D_l^2 + \int\sum_{i=1}\delta\boldsymbol{x}_i\cdot(d\boldsymbol{v}_i - \nabla_{\boldsymbol{x}}(\phi+\phi_{\text{ext}})dt) + \int\sum_{i=1}\delta\boldsymbol{v}_i\cdot(d\boldsymbol{x}_i(t) - \boldsymbol{v}_i(t)dt)\ , \tag{19}$$

leading to an explicit relation for the gradient of C with respect to $\boldsymbol{\alpha}$

$$\partial C/\partial\alpha_k = -\int\sum_{i=1}\delta\boldsymbol{x}_i\cdot\partial_{\boldsymbol{x}}(\partial_{\alpha_k}(\phi_{\text{ant}})dt = -\int dt\int d\boldsymbol{x}\left\{\delta\rho\ \partial_{\alpha_k}\phi_{\text{ext.}}\right\}. \tag{20}$$

Algorithm

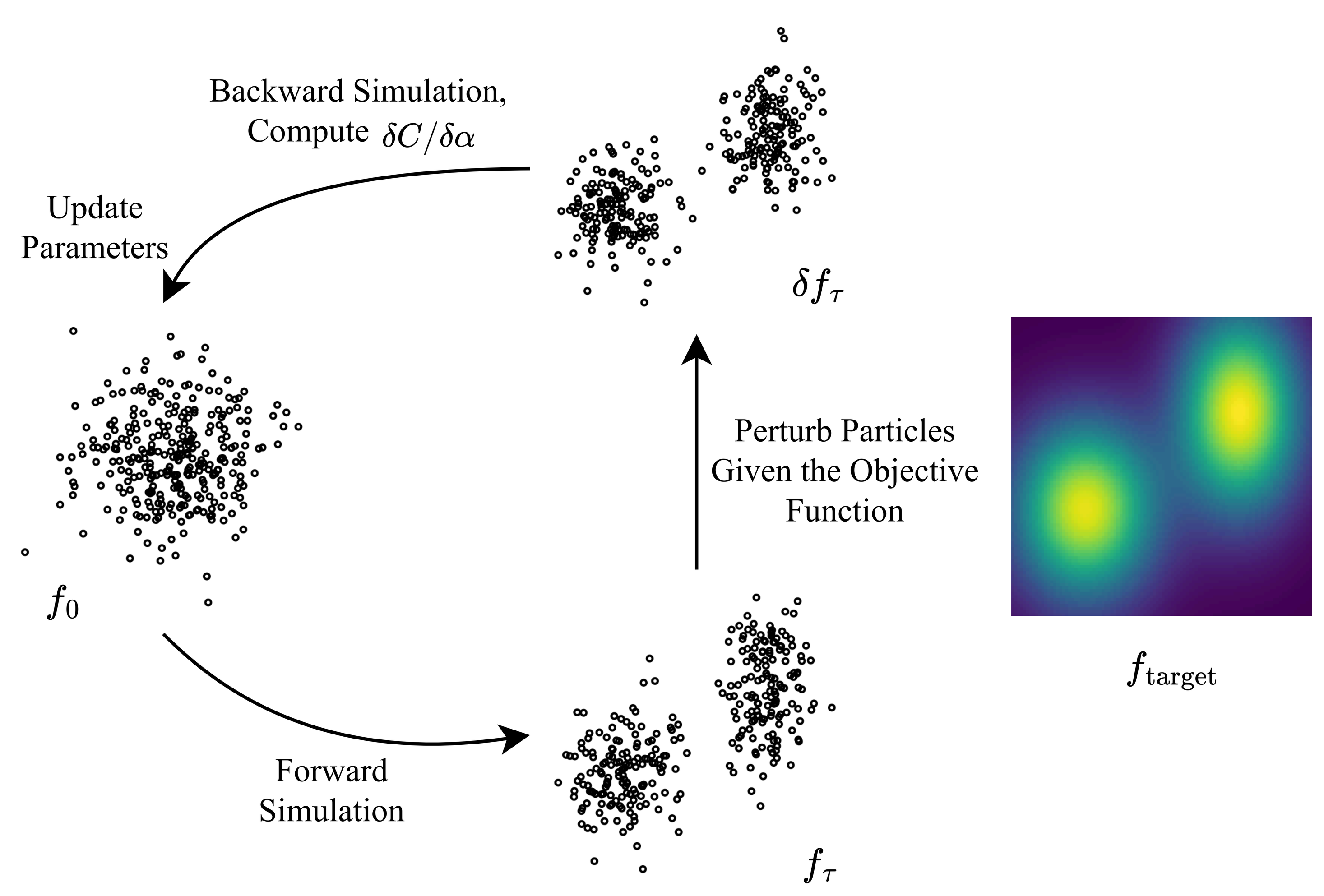

Figure 1: Algotithm of the adjoint method shown schematically

Next Steps

- Optimizing external force field for the Landau Damping test case in IPPL.
- Implementation in OPAL-X.