ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-TUP052

IMPLEMENTATION OF ADJOINT SENSITIVITY ANALYSIS IN WARP*

D. T. Abell†, V. B. Tembo, I. Haber, T. M. Antonsen, University of Maryland, College Park, MD, USA

Abstract

The design of accelerator lattices involves evaluating and optimizing Figures of Merit (FoMs) that characterize a beam's properties. These properties—hence the FoMs—depend on the many parameters that describe a lattice, including the strengths, locations, and possible misalignments of focusing elements. Often what is required is the gradient of the FoM with respect to each of the parameters. For systems that require numerical simulation, a naïve computation of a gradient requires one simulation for the "base case", plus one additional simulation for each parameter of interest—a daunting effort in the case of computationally demanding simulations with many parameters. Adjoint techniques allow one to extract gradient information from one base-case simulation plus an additional one or two carefully prepared simulations. We demonstrate these techniques using the accelerator simulation code WARP, and we present our proof-of-concept results using several different FoMs as the basis for adjoint analyses of a simple beamline with multiple parameters.

INTRODUCTION

The design of all accelerator systems today uses sophisticated, computationally intensive, numerical models. An essential element of the design process is a determination of how various performance metrics vary with respect to the many parameters that define an accelerator lattice. A determination of this dependence has at least two benefits: (i) It establishes the gradient of a given metric in parameter space, which facilitates design optimization using steepest descent. (ii) It yields the sensitivity of a particular design to errors in the lattice, and hence enables designers to set construction tolerances.

Adjoint methods provide a dramatically more efficient way to calculate parameter dependencies than straightforward parameter differencing, which requires a separate simulation run for each parameter being varied [1]. As an indication of the power of adjoint methods, we point out that adjoint techniques are used in the engineering community in a variety of contexts to compute sensitivities and gradients. The adjoint approach has previously been applied in circuit theory [2], electromagnetics [3], aerodynamics [4], plasma confinement [5–7], as well as in many other fields. They have been used to optimize the design of three-dimensional magnetohydrodynamic equilibria in stellarators [6, 7]. In the context of accelerators, we have recently deployed adjoint methods to optimize electrostatic [8] and magnetic [9] focusing systems, as well as to optimize the combination of quadrupole and solenoid magnetic fields needed to transform—in the presence of strong space charge—a flat beam to a round, angular-momentum-dominated beam [1].

A HAMILTONIAN ADJOINT RELATION

Lots of analysis and algebra (as sketched in the appendix) leads to the following *adjoint relation*, which involves two types of perturbations—true and adjoint, denoted respectively by superscripts (T) and (A):

$$\sum_j \frac{I_j}{I}\left(\delta p_j^{(T)}\cdot\delta x_j^{(A)} - \delta p_j^{(A)}\cdot\delta x_j^{(T)}\right)\Big|_0^L = \int_0^L \mathrm{d}z \int \mathrm{d}^2x_\perp \frac{q}{I}\Big\{ \left[-\delta\rho^{(A)}\,\delta\phi^{(T)} + \delta J^{(A)}\cdot\delta A^{(T)}\right] - \left[A \leftrightarrow T\right]\Big\}, \quad (1)$$

where j indexes particles in the beam, and $[A \leftrightarrow T]$ denotes the previous expression in square brackets with the superscripts (T) and (A) reversed. On the left-hand side (LHS), we have a net change in the symplectic area separating the trajectories of two different perturbed solutions. On the right-hand side (RHS), we see a space-time integral over mixed—(A) and (T)—products of charge and current density perturbations with, respectively, scalar and vector potential perturbations.

The utility of the adjoint relation derives from the freedom we have to assign the adjoint perturbations on the LHS: We shall assign them so as to make the LHS become the change in some relevant FoM. This constitutes the *difficult-to-compute* quantity of interest. The RHS we can evaluate using an additional pair of carefully instrumented simulations. This constitutes the *easy-to-compute* quantity that yields the change we seek to determine in the FoM of interest. Moreover, we can use the linearity of integration to accumulate individual contributions to the RHS, and compute thus—with just the one additional pair of simulations—the FoM *gradient* with respect to as many parameters as we deem relevant.

We focus on FoMs F that measure some aspect(s) of the particle phase-space distribution at the exit plane $z = L$. We write our F in the generic form

$$F = \sum_j \frac{I_j}{I} f(x_{\perp j}, p_{\perp j})\Big|_{z=L}. \quad (2)$$

To make the LHS of Eq. (1) look like a change in the FoM, we assume the True solution has no perturbations to the particle phase-space coördinates at the lattice entrance, $z = 0$, and we assume the Adjoint solution has particle perturbations at the lattice exit, $z = L$, given by the rules

$$\delta x_j^{(A)} = \frac{\partial f}{\partial p_j}\Delta, \quad \delta p_j^{(A)} = -\frac{\partial f}{\partial x_j}\Delta, \quad (3a)$$

* This work supported in part by the US Department of Energy, Office of Science, Office of High Energy Physics, under Grant #DE-SC-0025277.
† dabell@umd.edu

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-TUP052

where we define the *perturbation scale parameter*

$$\Delta = \Delta(\varepsilon) = \frac{x_{\text{typical}}\, p_{\text{typical}}}{F_{\text{typical}}}\varepsilon \tag{3b}$$

for small ε. Then, by the chain rule, the LHS becomes

$$\sum_j \frac{I_j}{I}\left[\delta p_{j\perp}^{(T)} \cdot \left(\frac{\partial f}{\partial p_{j\perp}}\Delta\right) + \left(\frac{\partial f}{\partial x_{j\perp}}\Delta\right) \cdot \delta x_{j\perp}{}^{(T)}\right]\Big|_L = \delta F \cdot \Delta. \tag{4}$$

To enforce the perturbation values at the exit plane, we must run the simulation forward, perturb the phase-space coördinates as in Eq. (3a), and then run the simulation in reverse.

To simplify the RHS of Eq. (1), we further constrain our perturbed solutions: The True solution we assume has perturbed beamline elements, corresponding to $\delta\phi^{(T)}$ and $\delta A^{(T)}$, whereas for the Adjoint solution we assume $\delta\phi^{(A)} = 0$ and $\delta A^{(A)} = 0$, which means the $[A \leftrightarrow T]$ term vanishes. In addition, the simple test we exhibit in the next section has only magnetic elements with purely transverse fields, hence only a longitudinal vector potential. In this case, the RHS further simplifies to the form

$$\begin{aligned}\frac{q}{I}\int_0^L \mathrm{d}z \int \mathrm{d}^2x_\perp\, \delta J_z{}^{(A)}\, \delta A_z{}^{(T)} \\ = \frac{q}{I}\int_0^L \mathrm{d}z \int \mathrm{d}^2x_\perp\, \delta J_z{}^{(A)} \sum_{l=1}^{N_e}\sum_{k=1}^{K_l} \frac{\partial A_z}{\partial g_{lk}}\delta g_{lk},\end{aligned} \tag{5}$$

where g_{lk} denotes a given parameter, with l labeling the N_e different elements, and k labeling each of the K_l variable parameters within element l. Putting together Eq. (4) and Eq. (5) and rearranging terms, we obtain the relative change in our FoM in the form

$$\frac{\delta F}{F} = \sum_{l=1}^{N_e}\sum_{k=1}^{K_l}\left(\frac{\delta g_{lk}}{g^0_{lk}}\right) \underbrace{\frac{g^0_{lk}}{2F\Delta}\int_0^L \mathrm{d}z\left[\chi^{(+)}_{lk}(z) - \chi^{(-)}_{lk}(z)\right]}_{\sigma_{lk}}, \tag{6a}$$

where

$$\chi^{(\pm)}(z) = \frac{q}{I}\int \mathrm{d}^2x_\perp\, J_z^{(\pm)}\, \frac{\partial A_z}{\partial g_{lk}}. \tag{6b}$$

Here (i) $J_z^{(\pm)}$ denotes the longitudinal current density with a perturbation of $\pm\Delta$ imposed; (ii) g^0_{lk} denotes either the design value, or some nominal scale value, of parameter g_{lk}; (iii) σ_{lk} denotes a *dimensionless sensitivity* of F with respect to changes in parameter g_{lk}; and (iv) the factor of 2 in the denominator of σ_{lk} appears because we are taking a symmetric difference.

Our goal is to compute the sensitivities σ_{lk}, and Eq. (6b) tells us that to do so, we need only (for the simple test below) instrument our simulations to compute the derivatives $\partial A_z/\partial g_{lk}$ and integrate the quantities $\chi^{(\pm)}(z)$.

A SIMPLE TEST AND RESULTS THEREOF

As as an initial, and very simple, test of the adjoint technique described above, we used Warp to simulate a 1 mA, 10 keV electron beam traversing a straight, 10 cm beamline containing a single, centrally-located, 5 cm, horizontally-focusing quadrupole. We implemented a smoothly-varying profile for the quadrupole, with its strength increasing as $\cos^2(\pi z/(2l_f))$ at the entrance, and decreasing in the same manner at the exit.

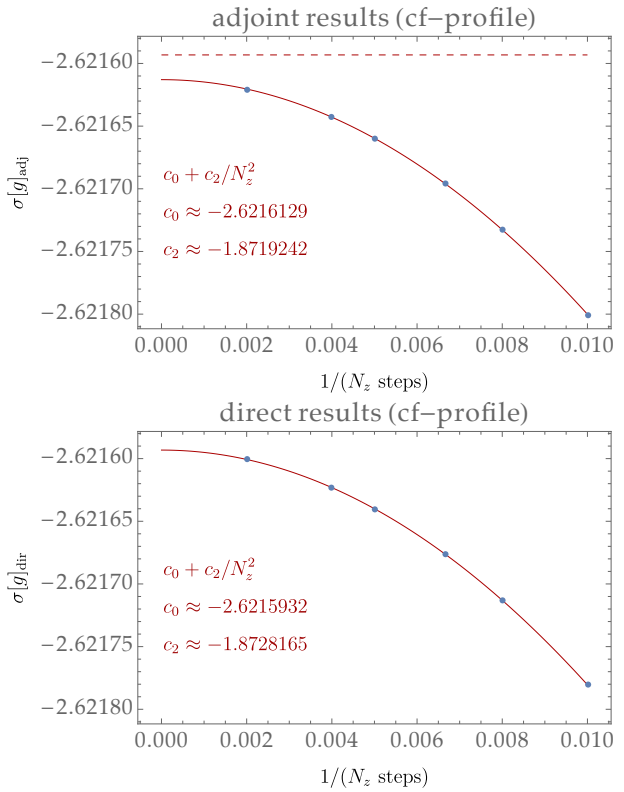

Figure 1: In the above graphics, the blue dots represent our adjoint (upper) and direct (lower) results for the sensitivity σ_g as a function of the reciprocal number of simulation steps. The inset expressions correspond to the quadratic fits shown.

For the FoM, we chose the horizontal rms beam size,

$$F = \sum_{j=1}^{N}\frac{I_j}{I}\frac{1}{2}x_j^2 = \frac{1}{2N}\sum_j x_j^2. \tag{7}$$

Here the index j runs over all N particles in the beam, and the second equality follows from the fact that we used equally-weighted particles in our simulations. We then ask for the sensitivity of F to the quadrupole gradient g about a nominal value $g^0 = -0.10\,\mathrm{T/m}$.

Anticipating some questions of sign, let us identify what we expect: If we increase g^0—*i.e.* $\delta g > 0$, making the gradient *less negative*—the quadrupole will focus the (negative) electron beam less strongly, and we therefore expect the horizontal beam size to increase, $\delta F > 0$. Since we seek

$$\sigma_g = \frac{\delta F/F^0}{\delta g/g^0} = \frac{\delta F/\delta g}{F^0/g^0}, \tag{8}$$

it follows that we expect a *negative* value.

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-TUP052

For values of ε in the range 0.001 to 0.2, we observed essentially no dependence of the computed sensitivity on ε, which implies these computations lie well within the linear regime. As shown in the upper graphic of Fig. 1, the computed sensitivity σ_g *does* vary with the simulation step-size, and converges quadratically as the step size vanishes. For comparison, we also computed the sensitivity σ_g directly from the slope of the curve (not shown) for our FoM plotted as a function of the quadrupole gradient. As shown in the lower graphic of Fig. 1, this direct (*i.e.* brute-force) computation of σ_g also converges quadratically as the step size vanishes. Quadratic fits in the two cases yield limiting values for σ_g (the constant terms, c_0) that agree to within 8 ppm.

CONCLUSION

Summary: In a very simple test, we obtained excellent (5-digit) agreement between adjoint and direct computations of a single sensitivity value. More generally, adjoint methods have the potential to reduce dramatically the computational burden of designing and optimizing complex accelerator lattices, with the ratio of adjoint to naïve effort being of order $(1+2) : (1+N_{\mathrm{param}})$. In addition, one can directly interpret the output values—the sensitivities σ_{lk}—in terms of a given parameter's impact on some relevant FoM. Moreover, whereas prior work developed adjoint methods on the basis of moment equations, here we started with particle equations of motion, which allows us to capture much more granular detail about the beam.

We emphasize that one can—with appropriately sophisticated simulations—include the effects of space-charge, beam loading, *etc.*, in the sensitivity computations described here.

Future Work: We are in the process of extending our current efforts to include multiple parameters in multiple beamline elements. We shall then implement the analysis of additional FoMs, with the intent to develop a *library of FoMs* for the evaluation of diverse types of machines, and to automate as much of this process as possible.

ACKNOWLEDGEMENTS

We express our particular gratitude to Mohsen Sadr and Andreas Adelmann of the Paul Scherrer Institute for early discussions on this subject.

APPENDIX

Here we *sketch* a proof of the adjoint relation Eq. (1) for the case of time-based codes.

To begin, we consider perturbations to both externally applied electromagnetic potentials and the particle trajectories. For perturbed potentials, we write $\phi + \delta\phi$ and $A + \delta A$, where $\delta\phi$ and δA can arise from changes in both self and external forces. And for perturbed trajectories, we write $\bar{x} = x + \delta x$ and $\bar{p} = p + \delta p$. Taken together, these perturbations modify the Hamiltonian, through first-order, to the form

$$\bar{H} = H + \frac{\partial H}{\partial x}\cdot\delta x + \frac{\partial H}{\partial p}\cdot\delta p + q(\delta\phi - v\cdot\delta A), \qquad (9)$$

where H here denotes the *unperturbed* Hamiltonian. Using the standard rules, we compute the corresponding equations of motion for the perturbed trajectories:

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = \frac{\partial H}{\partial p} + \frac{\partial^2 H}{\partial p\partial x}\cdot\delta x + \frac{\partial^2 H}{\partial p^2}\cdot\delta p - q\frac{\partial}{\partial p}(v\cdot\delta A),$$
$$\frac{\mathrm{d}\bar{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial x} - \frac{\partial^2 H}{\partial x^2}\cdot\delta x - \frac{\partial^2 H}{\partial x\partial p}\cdot\delta p - q\frac{\partial}{\partial x}(\delta\phi - v\cdot\delta A).$$

Subtract from this result the equations of motion for the unperturbed trajectories to obtain the following equations of motion for the trajectory perturbations:

$$\frac{\mathrm{d}\delta x_j}{\mathrm{d}t} = \frac{\partial^2 H}{\partial p\partial x}\cdot\delta x_j + \frac{\partial^2 H}{\partial p^2}\cdot\delta p_j - q_j\frac{\partial}{\partial p}(v_j\cdot\delta A), \qquad (10a)$$

$$\frac{\mathrm{d}\delta p_j}{\mathrm{d}t} = -\frac{\partial^2 H}{\partial x^2}\cdot\delta x_j - \frac{\partial^2 H}{\partial x\partial p}\cdot\delta p_j - q_j\frac{\partial}{\partial x}(\delta\phi - v_j\cdot\delta A). \qquad (10b)$$

Here we have added a subscript j to index individual particles in the beam. We have not added a subscript to the perturbations $\delta\phi$ and δA, but one must here evaluate them at the location of particle j.

If one dots Eq. (10a) with δp_j, dots Eq. (10b) with δx_j, and adds the results, one obtains an exact differential:

$$\frac{\mathrm{d}\delta x_j}{\mathrm{d}t}\cdot\delta p_j + \delta x_j\cdot\frac{\mathrm{d}\delta p_j}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\delta x_j\cdot\delta p_j). \qquad (11)$$

Now consider Eq. (11) with mixed perturbations. In particular, construct a symplectic area integrated over a time τ, and summed over (possibly weighted) particles j, according to

$$\sum_j \frac{I_j}{I}\left(\delta x_j^{(A)}\cdot\delta p_j^{(T)} - \delta x_j^{(T)}\cdot\delta p_j^{(A)}\right)\Big|_0^\tau. \qquad (12)$$

Replace the evaluation at 0 and τ by $\int_0^\tau \frac{\mathrm{d}}{\mathrm{d}t}$, carry out the time derivative, and insert Eq. (10). Keeping in mind that the potentials do not depend on momenta, one finds, after some tedious but routine effort, that all second-order derivatives of H cancel, with the result that

$$\sum_j \frac{I_j}{I}\left(\delta x_j^{(A)}\cdot\delta p_j^{(T)} - \delta x_j^{(T)}\cdot\delta p_j^{(A)}\right)\Big|_0^\tau = \sum_j q_j\int_0^\tau \mathrm{d}t\Big\{$$
$$\left(\delta x_j^{(A)}\cdot\frac{\partial}{\partial x} + \delta p_j^{(A)}\cdot\frac{\partial}{\partial p}\right)\left(\delta\phi^{(T)} - v_j\cdot\delta A^{(T)}\right) - (A\leftrightarrow T)\Big\},$$

where $(A\leftrightarrow T)$ indicates a corresponding set of terms with A and T interchanged.

From the Taylor series for a multi-variate function, one obtains the useful first-order relation $\delta u\cdot\partial f/\partial u \approx f(u+\delta u) - f(u)$. Using this relation together with the δ-function representation of charge density—$\rho(x) = \sum_j q_j\delta(x - x_j)$—one can rewrite the term that contains $\delta\phi^{(T)}$ in terms of an integral over $\delta\rho^{(A)}\,\delta\phi^{(T)}$. Similarly, one can rewrite the term that contains $\delta A^{(T)}$ in terms of an integral over $\delta J^{(A)}\cdot\delta A^{(T)}$.

The adjoint relation Eq. (1) results from a similar computation using the Hamiltonian with z as the independent coördinate.

ISBN: 978-3-95450-261-5 ISSN: 2673-7000 doi: 10.18429/JACoW-NAPAC2025-TUP052

REFERENCES

[1] L. Dovlatyan, B. L. Beaudoin, S. Bernal, I. Haber, D. F. Sutter, and T. M. Antonsen, "Optimization of flat to round transformers with self-fields using adjoint techniques", *Phys. Rev. Accel. Beams*, vol. 25, p. 044002, 2022. doi:10.1103/physrevaccelbeams.25.044002

[2] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities", *IEEE Trans. Circuit Theory*, vol. 16, pp. 318–323, 1969. doi:10.1109/TCT.1969.1082965

[3] N. K. Nikolova, R. Safian, E. A. Soliman, M. H. Bakr, and J. W. Bandler, "Accelerated gradient-based optimization using adjoint sensitivities", *IEEE Trans. Antennas Prop.*, vol. 52, pp. 2147–2157, 2004. doi:10.1109/TAP.2004.832313

[4] C. Othmer, "Adjoint methods for car aerodynamics", *J. Math. Industry*, vol. 4, p. 6, 2014. doi:10.1186/2190-5983-4-6

[5] T. M. Antonsen, and K. R. Chu, "Radio-frequency current generation by waves in toroidal geometry", *Phys. Fluids*, vol. 25, pp. 1295–1296, 1982. doi:10.1063/1.863906

[6] T. M. Antonsen, E. J. Paul, and M. Landreman, "Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria", *J. Plasma Phys.*, vol. 85, p. 905850207, 2019. doi:10.1017/S0022377819000254

[7] E. J. Paul, T. M. Antonsen, M. Landreman, and W. A. Cooper, "Adjoint approach to calculating shape gradients for three-dimensional magnetic confinement equilibria. Part 2. Applications", *J. Plasma Phys.*, vol. 86, p. 905860103, 2020. doi:10.1017/S0022377819000916

[8] T.M. Antonsen, D. Chernin, and J.J. Petillo, "Adjoint approach to beam optics sensitivity based on Hamiltonian particle dynamics", *Phys. Plasmas*, vol. 26, p. 013109, 2019. doi:10.1063/1.5079629

[9] T. M. Antonsen, B. L. Beaudoin, L. Dovlatyan, and I. Haber, "Adjoint approach to accelerator lattice design", in *Proc. NA-PAC'19*, Lansing, MI, USA, Sep. 2019, pp. 376–378. doi:10.18429/JACoW-NAPAC2019-TUPLM03