# Automatically find spring-mass-damper behavior to match observations using adjoint-based sensitivity: Part I Linear system

David A. Buchta*

*Coordinated Science Laboratory,*

*University of Illinois at Urbana–Champaign, IL 61820, USA*

* buchta1@illinois.edu

Adjoint-based sensitivity (see [1] for a review) is appealing for optimization of numerical simulations, and it has been used for data assimilation [2], control fluid flows, such as aeroacoustics [3–5], as well as, measuring parameter and field sensitivity [6, 7]. Similar methods have been used to control mechanical vibrations [8] and parameter identification [9]; here, efforts are concentrated on using a classical spring-mass-damper system (figure 1 a) to demonstrate automatically tuning parameters of a model to match observational data which is given figure 1 (b).

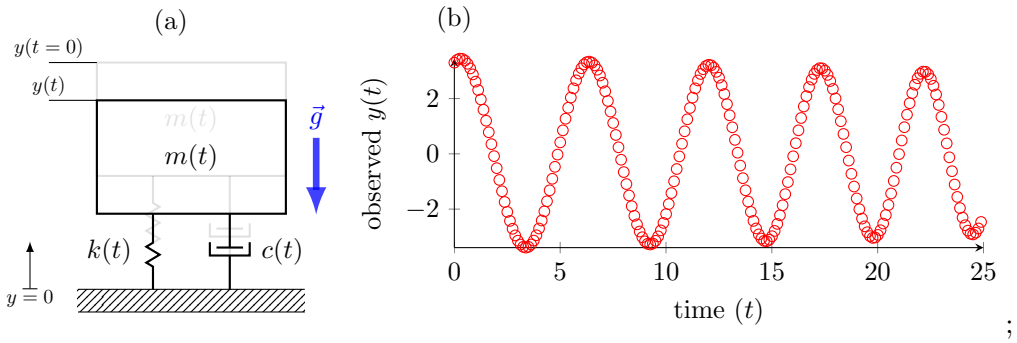


FIG. 1. (a) Schematic of spring-mass-damper system with components which are generally time dependent. (b) Observation data for the dynamics of height $y(t)$ for unknown device. Every 25 data points is shown for clarity.

Reducing data-mismatch with adjoint-directed model parameter adjustments is tested on time generic spring mass dampers (e.g. $m(t)$), which depending on the time discretization for their numerical solution, includes a parameter space up to size$(\vec{p}) > O(10^5)$ parameters. Despite significant degrees of freedom, forward and adjoint solutions, optimization, and post processing requires $\approx 1/60$ CPU hour using a 1.6 GHz Intel Core i5. The simulation code has been made publicly available [10]. The next section introduces the mathematical model applied to configuration in figure 1 and the formulation of the adjoint-based method used to identify system parameters. Section II provides results of the optimization and reveals the unknown device parameters which is compared against the optimized, adjoint-adjusted parameters. Finally, a summary is provided in section III.

## I. A VARIATIONAL APPROACH FOR OPTIMIZATION

This section explains the adjoint formulation used to automatically find the spring-mass-damper model to reduce the error between observation data ($y_{\text{obs}}$) and model output ($y$)

based on $\vec{p}$. We seek $\vec{p} = [m(t), k(t), c(t)]^T$ which minimizes

$$\mathcal{J} \equiv \int_{t=0}^{T} \mathcal{I}(y, \vec{p}; y_{\text{obs}}) dt = \int_{t=0}^{T} \left[y - y_{\text{obs}}\right]^2 dt. \tag{1}$$

Applying conservation of momentum to spring-mass-damper configuration in figure 1 (a), the dynamics follow

$$\mathcal{M}(y, \vec{p}) \equiv m(t)\frac{d^2}{dt^2}y + c(t)\frac{d}{dt}y + k(t)y = 0; \tag{2}$$

further given is the initial data $y = 3.3$ and $\frac{d}{dt}y = 1$. The initial value problem (2) is integrated in time using the standard second-order explicit Runge-Kutta method over the time of given data $0 \leq t \leq T$ for $T = 25$ [10].

The variation ($\delta$) of (2), multipled by adjoint field $y^\dagger$, is subtracted from the variation of the cost (1),

$$\int_{t=0}^{T} \delta\mathcal{I}(y, \vec{p})\, dt = \int_{t=0}^{T} \frac{\partial \mathcal{I}}{\partial \vec{p}} \delta\vec{p}\, dt + \int_{t=0}^{T} \frac{\partial \mathcal{I}}{\partial \vec{y}} \delta y\, dt \tag{3}$$

$$-\left( \int_{t=0}^{T} y^\dagger \delta\mathcal{M}(y, \vec{p})\, dt = \int_{t=0}^{T} y^\dagger \frac{\partial \mathcal{M}}{\partial y} \delta y\, dt + \int_{t=0}^{T} y^\dagger \frac{\partial \mathcal{M}}{\partial \vec{p}} \delta\vec{p}\, dt \right) \tag{4}$$

$$\delta\mathcal{J} = \int_{t=0}^{T} \left[\frac{\partial \mathcal{I}}{\partial y} - y^\dagger \frac{\partial \mathcal{M}}{\partial y}\right] \delta y\, dt + \int_{t=0}^{T} \left[\frac{\partial \mathcal{I}}{\partial \vec{p}} - y^\dagger \frac{\partial \mathcal{M}}{\partial \vec{p}}\right] \delta\vec{p}\, dt, \tag{5}$$

where $y^\dagger \delta\mathcal{M} = 0$ since $\mathcal{M} = 0$ by conservation of momentum (2).

With a careful choice of $y^\dagger$, the first bracketed term in (5) can be eliminated which provides the input ($\delta\vec{p}$) output ($\delta\mathcal{J}$) relationship we seek. To eliminate the problematic $\delta y$ term,

$$\int_{t=0}^{T} \left[\frac{\partial \mathcal{I}}{\partial y} - y^\dagger \frac{\partial \mathcal{M}}{\partial y}\right] \delta y\, dt,$$

$\mathcal{M}$ from (2) and cost from (1) are substituted. For the cost term, this yields

$$\int_{t=0}^{T} \frac{\partial \mathcal{I}}{\partial y} \delta y\, dt = \int_{t=0}^{T} 2[y - y_{\text{obs}}] \delta y\, dt \tag{6}$$

For the mass term, integration by parts is used to move time derivatives from $\delta y$ to $y^\dagger$,

$$\int_{t=0}^{T} y^\dagger m(t) \frac{d^2}{dt^2} \delta y\, dt = y^\dagger m(t) \frac{d\delta y}{dt}\bigg|_{t=0}^{T} - \int_{t=0}^{T} \frac{d\delta y}{dt} \frac{d}{dt} \left[y^\dagger m(t)\right] dt, \tag{7}$$

and applying integration by parts a second time to the right side of (7) yields

$$\int_{t=0}^{T} y^\dagger m(t) \frac{d^2}{dt^2} \delta y \, dt = y^\dagger m(t) \frac{d\delta y}{dt} \Big|_{t=0}^{T} \tag{8}$$

$$- \delta y \frac{d}{dt} \left[ y^\dagger m(t) \right] \Big|_{t=0}^{T} \tag{9}$$

$$+ \int_{t=0}^{T} \delta y \frac{d^2}{dt^2} \left[ y^\dagger m(t) \right] dt \tag{10}$$

The dissipation term is calculated in the same way by

$$\int_{t=0}^{T} y^\dagger c(t) \frac{d}{dt} \delta y \, dt = y^\dagger c(t) \delta y \Big|_{t=0}^{T} - \int_{t=0}^{T} \delta y \frac{d}{dt} \left[ c(t) y^\dagger \right] dt. \tag{11}$$

The spring constant term is simply

$$\int_{t=0}^{T} y^\dagger k(t) \delta y \, dt. \tag{12}$$

The adjoint system, by inspection of equations (6–12), is arranged as

$$m(t) \frac{d^2}{dt^2} y^\dagger + \left\{ 2 \frac{dm(t)}{dt} - c(t) \right\} \frac{d}{dt} y^\dagger + \left\{ \frac{d^2 m(t)}{dt^2} - \frac{dc(t)}{dt} + k(t) \right\} y^\dagger = 2 \left[ y(t) - y_{\text{obs.}} \right], \tag{13}$$

for which the left side is consistent with previous derivation [11]. To eliminate inhomogeneous data on $y(T)$, we choose $y^\dagger(t = T) = y_t^\dagger(t = T) = 0$ to eliminate $\delta y$ dependence at $t = T$, and since we are not considering $\delta y$ on the initial conditions, all time boundary terms in (8) and (11) at $t = 0$ are homogeneous. The adjoint system (13) is solved using the same methods to solve (2); the $y^\dagger$ is driven by the data mismatch appearing as a source term. Per its derivation, the solution $y^\dagger(t)$ zeros the $\delta y$-dependent term in (5), providing the instaneous gradient

$$\frac{\delta \mathcal{I}}{\delta \vec{p}} = -y^\dagger \frac{\partial \mathcal{M}}{\partial \vec{p}}, \tag{14}$$

for cost functionals independent of parameter $\vec{p}$ (e.g. a penalty-like parameter). Specifically, for the device considered here, these are

$$\frac{\delta \mathcal{I}}{\delta k} = -y^\dagger y, \tag{15}$$

$$\frac{\delta \mathcal{I}}{\delta m} = -y^\dagger \frac{d^2}{dt^2} y, \text{ and} \tag{16}$$

$$\frac{\delta \mathcal{I}}{\delta c} = -y^\dagger \frac{d}{dt} y, \tag{17}$$

which are used to reduce data mismatch in the following section. However, before embarking on optimization, it is prudent to check the derivation of equations (13) and (15–17) as well

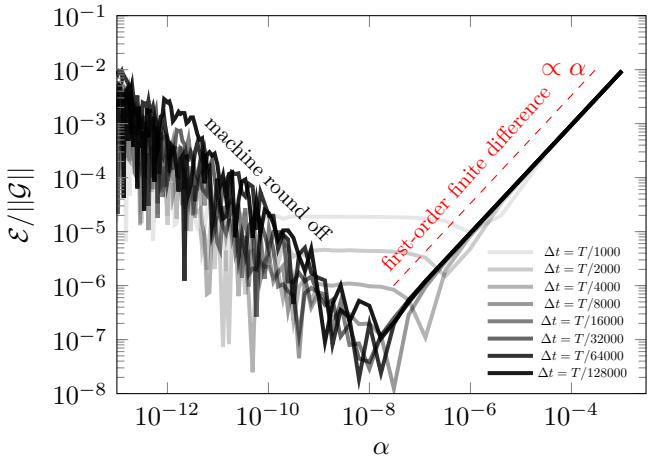


FIG. 2. Difference between adjoint-based gradient and first-order finite-difference approximation for different time resolutions ($\Delta t$).

as its numerical implementation. This is done by comparing the gradient to a first-order finite-difference approximation of it, following Taylor expansion

$$\mathcal{J}(y, \vec{p} + \delta\vec{p}) = \mathcal{J}(y_o, \vec{p}_o) + \left.\frac{\delta \mathcal{J}}{\delta \vec{p}}\right|_{y_o, \vec{p}_o} \delta\vec{p} + O(\delta\vec{p}^2). \tag{18}$$

The choice of parameter perturbation

$$\delta\vec{p} = -\alpha \underbrace{\left.\frac{\delta \mathcal{J}}{\delta \vec{p}}\right|_{y_o, \vec{p}_o}}_{\mathcal{G}}$$

with positive-valued scalar step-size $\alpha$ yields error measure

$$\mathcal{E} = \frac{\mathcal{J}(y, \vec{p} + \delta\vec{p}) - \mathcal{J}(y_o, \vec{p}_o)}{\alpha} + \mathcal{G}^2 + O(\alpha), \tag{19}$$

which is plotted in figure 2, and as expected, the error converges as first order $\propto \alpha$. For time resolution $\Delta t < T/16000$, the gradient from the adjoint is approximately machine precise. This verifies the implementation and derivation, to a degree, is accurate.

## II. RESULTS

Using the gradient from (15–17), the parameters, which are initially set to arbitrary values $\vec{p}_o = [1, 1, 1]^T$, are updated iteratively by

$$\vec{p}_i = \vec{p}_{i-1} - \alpha \partial \mathcal{J}_{\vec{p}} \tag{20}$$

so as to reduce the cost (1); details of the steepest descent optimization can be found in the demonstration software [10]. The route to optimal is demonstrated in figure 3. Iterations proceed until the cost and gradient, relative to their initial values, decrease below $10^{-4}$ as shown in figure 3 (a), which at $i = 10^3$ produces a sub-line width agreement to observed data in 3 (b). This indicates success, although more sophisticated optimization procedures may improve convergence.

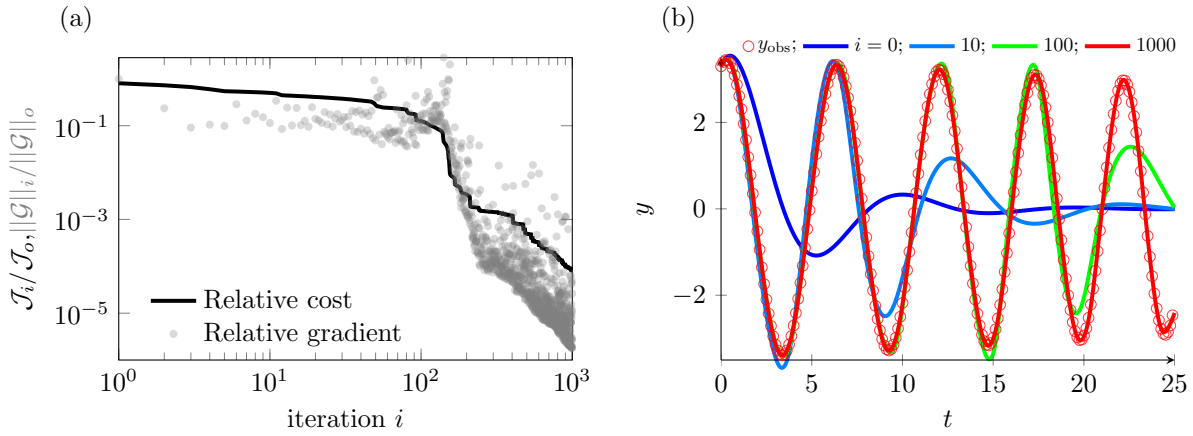


FIG. 3. (a) Changes in relative cost and gradient during optimization. (b) Comparison between observation data and automatically improving $(k, m, c)$ models. Iteration $i = 1000$ produces match to within sub-line width.

The optimized spring-mass-damper dynamics are shown in figure 4, and they are compared to the prior unknown device that generated data in figure 1 (b). For the observed data, the actual mass decayed linearly by 50% over the time horizon of interest ($T = 25$), while the spring and damper were constant values, $(k, c) = (1, 0)$. The optimized values however show stronger temporal dependence indicating the cost functional (in parameter space $\vec{p}$) has multiple optima, one of which was found by the adjoint-based optimization.

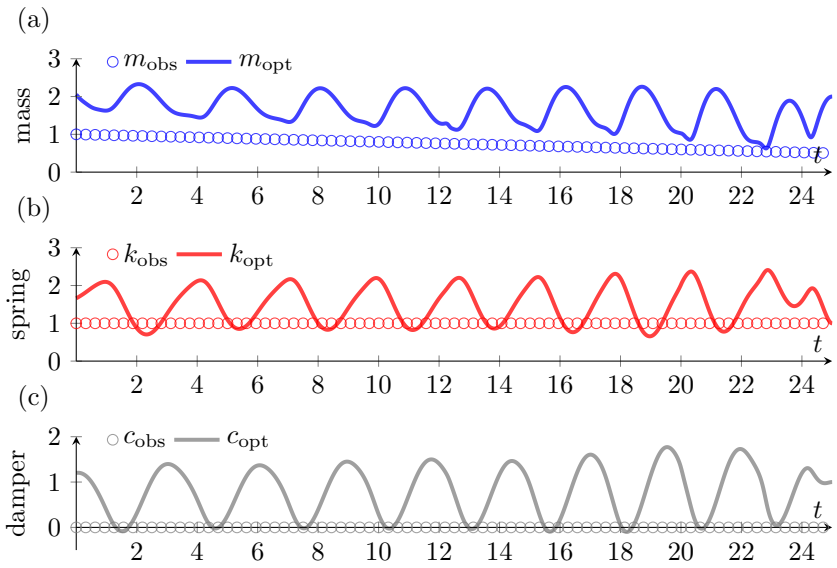


FIG. 4. Revealed (a) mass (b) spring (c) damper (symbols) generating observed data shown in figure 1 (b) compared with an optimized device (solid lines). For clarity, every twenty-five points are plotted for observed data.

## III. SUMMARY

Adjoint-based sensitivity and optimization for spring-mass-damper system was developed and applied to seek system parameters to reduce data mismatch. Automatic optimization was evaluated until the relative cost and gradient decreased below $10^{-4}$, which was deemed sufficient for identifying an optimum. Compared with the revelation of the device that generated the original data, the adjoint-based parameter adjustments can ultimately match the data, but the new system exhibits different $(k, m, c)$-behavior, suggesting the presence of multiple optima for this dynamical system. The computational cost of the adjoint-based optimization is independent of the number of parameters and can be extended to nonlinear systems and partial differential equations.

## IV. ACKNOWLEDGEMENTS

D. A. Buchta is grateful for the invitation from Professor Calvin Lui (Rose–Hulman Institute of Technology) to speak on this topic which spring boarded the development of

this demonstration, write-up, and Matlab code for undergraduate engineering students.

---


[1] R. E. Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503, 2006.

[2] R. H. Langland and N. L. Baker. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A: Dynamic Meteorology and Oceanography*, 56(3):189–201, 2004.

[3] J. B. Freund. Adjoint-based optimization for understanding and suppressing jet noise. *Journal of Sound and Vibration*, 330(17):4114–4122, 2011.

[4] J. Kim, D. J. Bodony, and J. B. Freund. Adjoint-based control of loud events in a turbulent jet. *Journal of Fluid Mechanics*, 741:28–59, 2 2014.

[5] D. Buchta, R. Vishnampet, D. J. Bodony, and J. B. Freund. A discrete adjoint-based shape optimization for shear-layer-noise reduction. In *22nd AIAA/CEAS Aeroacoustics Conference*.

[6] J. B. Freund, J. Kim, and R. H. Ewoldt. Field sensitivity of flow predictions to rheological parameters. *Journal of Non-Newtonian Fluid Mechanics*, 257:71 – 82, 2018.

[7] J. Capecelatro, D. J. Bodony, and J. B. Freund. Adjoint-based sensitivity and ignition threshold mapping in a turbulent mixing layer. *Combustion Theory and Modelling*, 23(1):147–179, 2019.

[8] C. Pappalardo and D. Guida. Use of the adjoint method for controlling the mechanical vibrations of nonlinear systems. *Machines*, 6(2):19, 2018.

[9] Thomas Lauß, Stefan Oberpeilsteiner, Wolfgang Steiner, and Karin Nachbagauer. The discrete adjoint method for parameter identification in multibody system dynamics. *Multibody system dynamics*, 42(4):397–410, 2018.

[10] Matlab-based adjoint solver (adjointexample-kmc), May 2019. URL `https://github.com/buchta1/adjointExample-kmc`.

[11] M. D. Greenberg. *Applications of Green's functions in science and engineering*. Courier Dover Publications, 2015.