The statistical mechanics of model space shuttles

Kris Innanen

ABSTRACT

Null space shuttles are an artifice which help quantify uncertainty in seismic full waveform inversion. Recent papers have framed these shuttles in terms of artificial dynamical systems, in which high-dimensional orbits are computed in a potential produced by the objective function. Here we pursue a natural extension of this, which is to develop a barometric equation for gases of these shuttles. This paper is largely mathematical, and ends with that equation. The main effort is in developing an approach based on the eigenvalues of the shuttle mass matrix and the Hessian matrix to produce a closed-form (if yet computationally expensive) expression for the partition function.

INTRODUCTION

The characterization of uncertainty in high-dimensional nonlinear inverse problems, especially those with expensive simulation steps, is very challenging, and it is fair to say that at present no single methodology is favoured. Each in one way or another requires that model space be explored in the environment of a constructed model vector. Methods include those which sample a posterior distribution based on prior and likelihood estimates, those which characterize curvature in the vicinity of a minimum with Hessian estimates, and those which move between points in model space with comparable misfit values, and others. Each step a method takes in exploring model space requires a simulation and evaluation of the objective function, which is where the cost is incurred.

The third of the classes of methods discussed above, which involves *null space shuttles*, appear to be promising for full waveform inversion because of the balance these methods create between number of simulations and volume of model space explored. Keating and Innanen (2021) limit the exploration of model space to examining those points which maintain a comparable objective function value, and also maximally violate a selected hypothesis about the geology of the medium. This greatly reduces the size of the uncertainty characterization in a way that also supports geological interpreters making use of the FWI results. Fichtner et al. (2021) set the problem up as an artificial problem in classical mechanics, in which a position in model space (i.e., a constructed model vector) is treated as a dynamic quantity, with an artificial mass and momentum, which orbits model space subject to an artificial gravitational potential set equal to the objective function. Initializing the orbit with a particular momentum, then assuming that the dynamical behaviour to follow occurs based on the Hamiltonian (e.g., Goldstein et al., 2002) induced by the artifices above, produces a path taken by the model point that explores regions of space with similar objective function values. This trades away much of the savings in simulations of Keating's approach, in order to explore a much wider swath of allowable models.

Although it threatens to expand the volume of model space to be discussed, the current work of Fichtner's group appears to call for a statistical-mechanical treatment in addition to their "orbital dynamics" treatment. The purpose of this paper is to formulate and examine the mathematical features of such a treatment. In this approach, we envision not a single "space shuttle" orbiting model space, but instead a gas of shuttles*. This gas will, at least initially, not be one of null-space shuttles, because it will not be constrained to a single value or infinitesimal range of values. So, we will refer to it as a gas of model space shuttles instead.

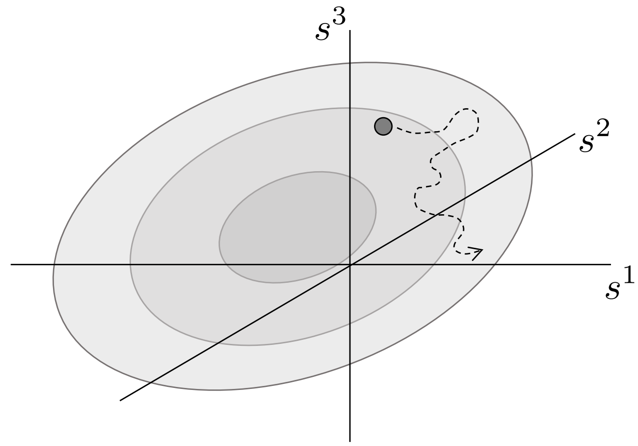

FIG. 1. A gas of model-space shuttles.

The mathematics should call to mind a classical result of statistical mechanics called the law of atmospheres, or the barometric law. The barometric law predicts the variation in elevation of the pressure and density of the atmosphere, based on an application of the Maxwell-Boltzmann statistics, and a statement of the gravitational potential energy carried by an individual air molecule. Rough approximations of macroscopic features of the atmosphere are determined in this law from microscopic equations of a single molecule, and simple statistical assumptions about the likelihood of the molecule finding itself in a particular state of kinetic and potential energy. The gas of model space shuttles likewise exerts an artificial pressure, which varies in model space in relation to the minimum of the objective function (Figure 1). Perhaps more importantly, the system produces an analytically tractable partition function, in terms of the determinants of the Hessian and the mass matrix and their eigenvalues.

MODEL SPACE SHUTTLES

A model is an array of unknowns forming a vector s^μ in a space S of dimension N. In the shuttle space approach, S is treated as a physical space through which fictitious objects (shuttles) can move about as a time-like parameter τ evolves. A model, as we typically understand the term, is in this approach a position that can be occupied by a shuttle. In order for the shuttle to move between different model positions, it also has momentum, a vector p^μ also with dimension N, and in order to describe it dynamically, it also has inertia, which appears in the form of a symmetric posititive-definite matrix $m_{\mu\nu}$. The Hamiltonian

*Technically, we will set up a gas of a single shuttle, which in our approach will have statistical properties that easily generalize to a gas of many shuttles.

of this shuttle system is defined to be

$$H(p, s) = K(p) + U(s), \tag{1}$$

where $K(p)$ is the kinetic energy of the shuttle and $U(s)$ is the potential energy. The kinetic energy is determined by the mass matrix and the shuttle momentum via

$$K(p) = \frac{1}{2} p^\mu m^{-1}_{\mu\nu} p^\nu. \tag{2}$$

This is a generalization in the sense that, if $m_{\mu\nu}$ is simplified such that all off diagonal elements are zero and the diagonals are equal, i.e., $m_{\mu\nu} = m\delta_{\mu\nu}$, then K reduces to the recognizable form

$$K(p) = \frac{(p^1)^2}{2m} + ... + \frac{(p^N)^2}{2m}. \tag{3}$$

In the shuttle system, an artificial potential energy is provided by the objective function. We adopt a linearized forward model in which the ith predicted datum is given by

$$d^i_p = F^i_\mu s^\mu, \tag{4}$$

such that if d^i_o is the ith observed datum the residual is

$$r^i = F^i_\mu s^\mu - d^i_o. \tag{5}$$

Using these quantities, we define an objective function / potential energy

$$U(s) = \frac{1}{2} r^i(s) r_i(s) + \frac{1}{2} s^\mu w_{\mu\nu} s^\nu, \tag{6}$$

where $w_{\mu\nu} = k^\lambda_\mu k_{\lambda\nu}$, and k^λ_μ is a roughening operator enhancing un-wanted model components (Harlan, 1994). This $U(s)$ can be re-written in the convenient form

$$U(s) = \frac{1}{2} s^\mu \phi_{\mu\nu} s^\nu - \varphi_\lambda s^\lambda + \psi, \tag{7}$$

where

$$\phi_{\mu\nu} = F^i_\mu F_{i\nu} + w_{\mu\nu}, \;\; \varphi_\lambda = F^i_\lambda d_{io}, \;\; \text{and} \;\; \psi = \frac{1}{2} d^i_o d_{io}. \tag{8}$$

In total, therefore, the Hamiltonian of the shuttle system is

$$H(p, s) = \frac{1}{2} p^\mu m^{-1}_{\mu\nu} p^\nu + \frac{1}{2} s^\mu \phi_{\mu\nu} s^\nu - \varphi_\lambda s^\lambda + \psi. \tag{9}$$

For linear problems per (4), and for non-linear problems that are approximately linear near the solution, the matrix ϕ is symmetric and positive-definite. We will also assume this of the mass matrix in the developments below.

Null space shuttles

The null-space shuttle approach is then implemented by setting starting conditions s_0^μ and p_0^μ such that regions of model space with comparable misfits are explored by solving the Hamiltonian equations for trajectories $s^\mu(\tau)$:

$$\frac{\partial s^\mu}{\partial \tau} = g^{\mu\nu}\frac{\partial H}{\partial p^\nu}, \text{ and } \frac{\partial p^\mu}{\partial \tau} = -g^{\mu\nu}\frac{\partial H}{\partial s^\nu}. \tag{10}$$

SHUTTLE GAS STATISTICS

Here we instead use the set-up as the starting point for a statistical analysis of the shuttles. This involves solving for the Boltzmann distribution for a gas of one model space shuttle in equilibrium with an environment at temperature $T = 1/\beta$:

$$P(p,s) = \frac{1}{Z(\beta)}e^{-\beta H(p,s)}, \tag{11}$$

where the partition function is

$$Z(\beta) = \int_{-\infty}^{\infty} dp^1 ... \int_{-\infty}^{\infty} dp^N \int ds^1 ... \int ds^N e^{-\beta H(p,s)}, \tag{12}$$

and the integrals over the components of s can be chosen to be over some well-defined bounds, such that $V = \int ds^1 ... \int ds^N$ is a finite volume of the space, or with one or all of the limits being at infinity, so that distributions which decay exponentially can be more easily modelled.

Shuttles in a potential

A gas consisting of a single shuttle with the Hamiltonian in (10) has the partition function

$$Z(\beta) = e^{-\beta\psi} I_P I_S, \tag{13}$$

where

$$I_P = \int_{-\infty}^{\infty} dp^1 ... \int_{-\infty}^{\infty} dp^N e^{-\frac{\beta}{2}\left(p^\mu m^{-1}_{\mu\nu} p^\nu\right)} \tag{14}$$

and, choosing the option of an unbounded model space,

$$I_S = \int_{-\infty}^{\infty} ds^1 ... \int_{-\infty}^{\infty} ds^N e^{-\frac{\beta}{2}\left(s^\mu \phi_{\mu\nu} s^\nu - 2\phi_\lambda s^\lambda\right)}. \tag{15}$$

The integrals in (14)-(15) can be difficult to evaluate when the ϕ and m matrices have off-diagonal elements. However, if we let $s^\mu = u^\mu_\nu x^\nu$ and $p^\mu = v^\mu_\nu y^\nu$, where u and v are matrices containing the eigenvectors of $\phi_{\mu\nu}$ and $m_{\mu\nu}$ respectively, the quadratic forms simplify to

$$p^\mu m^{-1}_{\mu\nu} p^\nu = \frac{(y^1)^2}{M_1} + ... + \frac{(y^N)^2}{M_N}, \tag{16}$$

where M_n is the nth eigenvalue of the mass matrix $m_{\mu\nu}$, and

$$s^\mu \phi_{\mu\nu} s^\nu = \Phi_1 (x^1)^2 + ... + \Phi_N (x^N)^2, \tag{17}$$

where Φ_n is the nth eigenvalue of the Hessian matrix $\phi_{\mu\nu}$. The product of integrals in (14) remains unchanged after the transformation, so

$$I_P = \left(\int_{-\infty}^{\infty} dy^1 e^{-\frac{\beta}{2M_1}(y^1)^2} \right) \times ... \times \left(\int_{-\infty}^{\infty} dy^N e^{-\frac{\beta}{2M_N}(y^N)^2} \right) = \left(\frac{2\pi}{\beta} \right)^{N/2} \sqrt{\det m},$$

using $\int_{-\infty}^{\infty} dq e^{-aq^2} = \sqrt{\pi/a}$ and $\det m = M_1 \times ... \times M_N$. The model space integrals are similarly uncoupled:

$$I_S = \left\{ \int_{-\infty}^{\infty} dx^1 \exp \left[- \left(\frac{\beta}{2} \Phi_1 \right) (x^1)^2 + \beta Z_1 x^1 \right] \right\} \times ..., \tag{18}$$

where Z_1 is the first component of $Z_\mu = u_\mu^\nu \varphi_\nu$, hence

$$I_S = \left(\frac{2\pi}{\beta} \right)^{N/2} \sqrt{\frac{1}{\det \phi}} \exp \chi, \tag{19}$$

using $\int_{-\infty}^{\infty} e^{-ax^2 - bx} = \sqrt{\pi/a}\, e^{b^2/4a}$ for each of the N integrals, and defining

$$\chi = \frac{1}{2} \sum_{k=1}^{N} \frac{Z_k^2}{\Phi_k}. \tag{20}$$

The partition function is then

$$Z(\beta) = \left(\frac{2\pi}{\beta} \right)^N \sqrt{\frac{\det m}{\det \phi}} \exp \big[- \beta(\psi - \chi) \big]. \tag{21}$$

The Boltzmann distribution for the single shuttle gas is therefore

$$P(p, s) = \left(\frac{2\pi}{\beta} \right)^{-N} \sqrt{\frac{\det \phi}{\det m}} \exp \left[-\beta \left(\frac{1}{2} p^\mu m_{\mu\nu}^{-1} p^\nu + \frac{1}{2} s^\mu \phi_{\mu\nu} s^\nu - \varphi_\lambda s^\lambda + \chi \right) \right]. \tag{22}$$

The number density $n(s)$ of N independent shuttles obeying these statistics is N times the integral of $P(p, s)$ over the momentum coordinates:

$$n(s) = N \left(\frac{2\pi}{\beta} \right)^{-N/2} \sqrt{\det \phi} \exp \left[-\beta \left(\frac{1}{2} s^\mu \phi_{\mu\nu} s^\nu - \varphi_\lambda s^\lambda + \chi \right) \right]. \tag{23}$$

This in fact allows us to define a pressure exerted by the shuttle gas via $p = nT$ (letting un-indexed lower-case p denote pressure), or

$$p(s) = NT \, (2\pi T)^{-N/2} \sqrt{\det \phi} \exp \left[- \left(\frac{1}{2} s^\mu \phi_{\mu\nu} s^\nu - \varphi_\lambda s^\lambda + \chi \right) / T \right]. \tag{24}$$

ACKNOWLEDGEMENTS

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19, and in part by an NSERC-DG.

REFERENCES

Fichtner, A., Zunino, A., Gebraad, L., and Boehm, C., 2021, Autotuning hamiltonian monte carlo for efficient generalized nullspace exploration: Geophysical Journal International, **227**, 941–968.

Goldstein, H., Safko, J. L., and Poole, C. P., 2002, Classical Mechanics: Addison-Wesley, 3rd edn.

Harlan, W. S., 1994, Regularization by model reparameterization: http:// billharlan. com/ pub/ papers/ regularization/ regularization.html.

Keating, S. D., and Innanen, K. A., 2021, Null-space shuttles for targeted uncertainty analysis in full-waveform inversion: Geophysics, **86**, No. 1.