SeisAcoustic

*Wenlei Gao**

ABSTRACT

Acoustic wave field modelling and inversion

METHOD

Our derivation start from the first-order acoustic wave equation which support variable density

$$
\begin{aligned}
\frac{\partial v_z}{\partial t} &= b\frac{\partial p}{\partial z} \\
\frac{\partial v_x}{\partial t} &= b\frac{\partial p}{\partial x} \\
\frac{\partial p}{\partial t} &= k(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x}) + f(t)
\end{aligned}
\qquad (1)
$$

where b is buoyancy, the reciprocal of density ρ, k is bulk modulus ($k = \rho v^2$), v_z and v_x are vertical and horizontal particle velocities and p is pressure. $f(t)$ is additive source term. To eliminate undesirable boundary reflections, we implement PML absorbing boundary conditions. The pressure component p is split into two unphysical components p_z and p_x, where $p = p_z + p_x$. Equation 1 can be reformulated as

$$
\begin{aligned}
\frac{\partial v_z}{\partial t} + r_z\,v_z &= b\frac{\partial\,(p_z + p_x)}{\partial z} \\
\frac{\partial v_x}{\partial t} + r_x\,v_x &= b\frac{\partial\,(p_z + p_x)}{\partial x} \\
\frac{\partial p_z}{\partial t} + r_z\,p_z &= k\frac{\partial v_z}{\partial z} \\
\frac{\partial p_x}{\partial t} + r_x\,p_x &= k\frac{\partial v_x}{\partial x} + f(t)
\end{aligned}
\qquad (2)
$$

As the source term is additive, it can be added to either one of the last two equations. γ_z and γ_x are the PML damping coefficients in vertical and horizontal directions, respectively. The bulk modulus is set to 0 when free surface boundary condition is implemented for the top boundary.

Discretization of Acoustic wave equation

We use stagger grid finite difference method to discretize equation 2, the particle v_z is put on the stagger grid $[i + \frac{1}{2}, j, it]$, horizontal particle velocity v_x is put on the stagger grid

$[i, j+\frac{1}{2}, k]$, and the pressure component is put on the grid $[i, j, it+\frac{1}{2}]$. Where i, j, it are the grid index in vertical, horizontal, time directions, respectively. Figure 1 shows the locations of wave field components and model parameters, respectively.
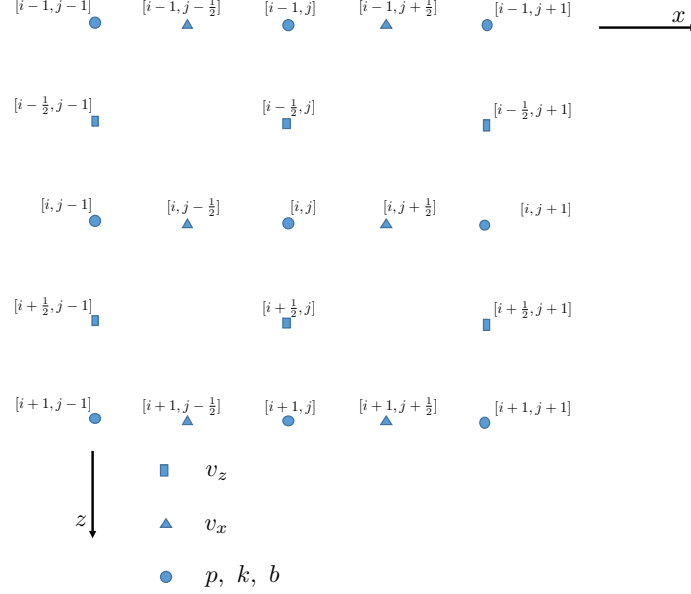

Figure 1: Stagger grid

$\mathbf{v}_z$ time stepping

$$\frac{\partial v_z}{\partial t} + \gamma_z \, v_z = b\frac{\partial p}{\partial z}. \tag{3}$$

The time partial derivative operator $\frac{\partial}{\partial t}$ is approximated by second-order central finite difference and spatial partial derivative operator is approximated by $2N^{th}$-order central finite difference. The center of both side of the equation is on $[i + \frac{1}{2}, j, it + \frac{1}{2}]$.

$$\frac{v_z[i + \frac{1}{2}, j, it + 1] - v_z[i + \frac{1}{2}, j, it]}{dt} + \gamma_z[i + \frac{1}{2}]\frac{v_z[i + \frac{1}{2}, j, it + 1] + v_z[i + \frac{1}{2}, j, it]}{2}$$
$$= b[i + \frac{1}{2}, j]\left(\frac{1}{dz}\sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}]\right)\right) \tag{4}$$

where dt is the time step size and dz is the grid size in vertical direction. $b[i + \frac{1}{2}, j]$ is approximated by averaging the buoyancy on neighbouring two grid points, so $b[i + \frac{1}{2}, j] = (b[i, j] + b[i+1, j])/2$. a_n is the finite difference coefficients, which can be obtained by Taylor-series expansion. p is the summation of p_z and p_x. We times both side of the equation 4 by $2dt$, we get

$$2(v_z[i + \frac{1}{2}, j, it + 1] - v_z[i + \frac{1}{2}, j, it]) + dt\,\gamma_z[i + \frac{1}{2}](v_z[i + \frac{1}{2}, j, it + 1] + v_z[i + \frac{1}{2}, j, it])$$
$$= 2b[i + \frac{1}{2}, j]\,dt\left(\frac{1}{dz}\sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}]\right)\right). \tag{5}$$

we move all the terms except $v_z[i + \frac{1}{2}, j, it + 1]$ to the right hand side of equation 5, we get

$$
\begin{aligned}
(2 + dt\,\gamma_z[i + \frac{1}{2}])v_z[i + \frac{1}{2}, j, it + 1] &= (2 - dt\,\gamma_z[i + \frac{1}{2}])v_z[i + \frac{1}{2}, j, it] \\
&+ 2b[i + \frac{1}{2}, j]\,dt \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}] \right) \right).
\end{aligned}
\tag{6}
$$

Both sides are divided by $2 + dt\,\gamma_z[i + \frac{1}{2}]$, we get the final form used for updating v_z

$$
\begin{aligned}
v_z[i + \frac{1}{2}, j, it + 1] &= \frac{2 - dt\,\gamma_z[i + \frac{1}{2}]}{2 + dt\,\gamma_z[i + \frac{1}{2}]} v_z[i + \frac{1}{2}, j, it] \\
&+ \frac{2b[i + \frac{1}{2}, j]\,dt}{2 + dt\,\gamma_z[i + \frac{1}{2}]} \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}] \right) \right)
\end{aligned}
\tag{7}
$$

To facilitate formulating the forward modelling into matrix-vector format, we vectorize the vertical particle velocity component v_z into a long vector $\mathbf{v}_z$ length of $N_z \cdot N_x \times 1$. Supposing the size of modelling area is $n_z \times n_x$ grid points, and the number of PML layers is $npml$. So

$$
N_z = \begin{cases} n_z + 2 \cdot npml, & \text{if } free\_surface = false \\ n_z + npml, & \text{if } free\_surface = true \end{cases}
\tag{8}
$$

The of N_z depends on the top boundary condition we implemented for the wave equation. There are two options, free surface boundary condition or absorbing boundary condition. We also give the width of the model after padding PML layers

$$
N_x = n_x + 2 \cdot npml.
\tag{9}
$$

we create a diagonal matrix $\mathbf{M}_{vz}\mathbf{B}_{vz}$, size of $N_z \cdot N_x \times N_z \cdot N_x$, whose diagonal elements saves the result of $\frac{2 - dt\,\gamma_z[i + \frac{1}{2}]}{2 + dt\,\gamma_z[i + \frac{1}{2}]}$. We also pre-compute an another diagonal matrix, still size of $N_z \cdot N_x \times N_z \cdot N_x$, called $\mathbf{M}_{vz}\mathbf{B}_p$, saves the result of $\frac{2b[i + \frac{1}{2}, j]\,dt}{2 + dt\,\gamma_z[i + \frac{1}{2}]}$. As column-wise lexicographical order is used for matrix vectorization, so l^{th} diagonal element is related to the sub-index $[i, j]$ via the relationship

$$
l = (j - 1) \times N_z + i
$$

We construct a sparse square matrix $\mathbf{d}_p\mathbf{d}_z$ size of $N_z \cdot N_x \times N_z \cdot N_x$ to represent the vertical partial derivative operator $\frac{\partial}{\partial z}$. In another words, the computation of

$$
\left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}] \right) \right)
$$

is realized by perform a sparse matrix-vector multiplication $\mathbf{d}_p\mathbf{d}_z \cdot \mathbf{p}$. we use an toy example to explain the construction of $\mathbf{d}_p\mathbf{d}_z$ for a 1D model which includes 5 grid points, and the partial derivative operator is approximated by 4th order finite-difference. So the vertical partial derivative can be expressed as

$$
\begin{bmatrix} \frac{\partial p}{\partial z}[1 + \frac{1}{2}] \\ \frac{\partial p}{\partial z}[2 + \frac{1}{2}] \\ \frac{\partial p}{\partial z}[3 + \frac{1}{2}] \\ \frac{\partial p}{\partial z}[4 + \frac{1}{2}] \\ \frac{\partial p}{\partial z}[5 + \frac{1}{2}] \end{bmatrix} = \frac{1}{dz} \begin{bmatrix} -a_1 & a_1 & a_2 & & \\ -a_2 & -a_1 & a_1 & a_2 & \\ & -a_2 & -a_1 & a_1 & a_2 \\ & & -a_2 & -a_1 & a_1 \\ & & & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} p[1] \\ p[2] \\ p[3] \\ p[4] \\ p[5] \end{bmatrix}
\tag{10}
$$

We set the matrix

$$\mathbf{d}_p\mathbf{d}_z = \frac{1}{dz} \begin{bmatrix} -a_1 & a_1 & a_2 & & \\ -a_2 & -a_1 & a_1 & a_2 & \\ & -a_2 & -a_1 & a_1 & a_2 \\ & & -a_2 & -a_1 & a_1 \\ & & & -a_2 & -a_1 \end{bmatrix}, \tag{11}$$

In the above formulations, we assume the value of wave field outside the computation area is 0. Equation 10 can be simplified as

$$\frac{\partial \mathbf{p}}{\partial z} = \mathbf{d}_p\mathbf{d}_z \cdot \mathbf{p} \tag{12}$$

It is straight-forward to extend the above formulation to 2D cases. We just need to change the definition of $\mathbf{d}_p\mathbf{d}_z$ as

$$\mathbf{d}_p\mathbf{d}_z = \mathbf{I}_{N_x} \otimes \mathbf{d}_p\mathbf{d}_z, \tag{13}$$

where $\mathbf{I}_{N_x}$ is an identity matrix size of $N_x \times N_x$ and $\mathbf{d}_p\mathbf{d}_z$ is sparse matrix defined in equation 13, $\otimes$ represent Kronecker product (tensor product). With these definitions, the one-step updating equation for $\mathbf{v}_z$ can be written as

$$\mathbf{v}_z^{it+1} = \mathbf{M}_{vz}\mathbf{B}_{vz} \cdot \mathbf{v}_z^{it} + \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z \cdot \mathbf{p}^{it+\frac{1}{2}} \tag{14}$$

Note: In the real implementations, we don't have to save the $\mathbf{M}_{vz}\mathbf{B}_{vz}$ as a diagonal matrix. We just save it as a vector, the multiplication between the diagonal matrix and the vector $\mathbf{v}_z$ is equivalent to the element-wise multiplication between two vectors of same length. It is not hard to discover that the elements of $\mathbf{M}_{vz}\mathbf{B}_{vz}$ are repeated column-by-column from equation 7. To further save the memory cost, the value of $\mathbf{M}_{vz}\mathbf{B}_{vz}$ can be saved as a vector length of N_z. Similarly, we can compute the vertical partial derivative of v_z column-by-column, it means that we don't need to perform Kronecker product in creating $\mathbf{d}_p\mathbf{d}_z$, which is savd as a $N_z \times N_z$ sparse matrix in the code.

Outside of PML absorbing boundary area, the damping coefficients γ_z is 0, so the equation 7 is reduced to

$$\begin{aligned} v_z[i + \frac{1}{2}, j, it + 1] &= v_z[i + \frac{1}{2}, j, it] \\ &+ b[i + \frac{1}{2}, j] \, dt \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(p[i + n, j, it + \frac{1}{2}] - p[i - n + 1, j, it + \frac{1}{2}] \right) \right) \end{aligned} \tag{15}$$

This formulation is used for wave field reconstruction backward or forward when the boundary value and the wave field at the last time step is savd. As the size the model is shrinked when the PML layers are trimed, we need to create another diagonal matrix called $\mathbf{R}_{vz}\mathbf{B}p$, whose diagonal elements save the result of $b[i + \frac{1}{2}, j] \, dt$. The size of the diagonal matrix $\mathbf{R}_{vz}\mathbf{B}_p$ is $n_z \cdot n_x \times n_z \cdot n_x$. An new sparse matrix $\mathbf{r}_p\mathbf{d}_z$ is created for approximating the vertical partial derivative of v_z outside of PML boundary layers. The size of $\mathbf{r}_p\mathbf{d}_z$ is also $n_z \cdot n_x \times n_z \cdot n_x$. With these defined sparse matrix, equation 15 can be written into the vectorized form as

$$\mathbf{v}_z^{it+1} = \mathbf{v}_z^{it} + \mathbf{R}_{vz}\mathbf{B}_p \cdot \mathbf{r}_p\mathbf{d}_z \cdot \mathbf{p}^{it+\frac{1}{2}} \tag{16}$$

We also need the adjoint operator of the vertical partial derivative operator when computing the adjoint wave field, the adjoint operator of $\mathbf{d}_p\mathbf{d}_z$ is given as

$$
\begin{bmatrix} p[1] \\ p[2] \\ p[3] \\ p[4] \\ p[5] \end{bmatrix} = -\frac{1}{dz}
\begin{bmatrix}
a_1 & a_2 & & & \\
-a_1 & a_1 & a_2 & & \\
-a_2 & -a_1 & a_1 & a_2 & \\
& -a_2 & -a_1 & a_1 & a_2 \\
& & -a_2 & -a_1 & a_1
\end{bmatrix}
\begin{bmatrix}
\frac{\partial p}{\partial z}[1+\frac{1}{2}] \\
\frac{\partial p}{\partial z}[2+\frac{1}{2}] \\
\frac{\partial p}{\partial z}[3+\frac{1}{2}] \\
\frac{\partial p}{\partial z}[4+\frac{1}{2}] \\
\frac{\partial p}{\partial z}[5+\frac{1}{2}]
\end{bmatrix}
$$

$\mathbf{v}_x$ time stepping

The equation related to horizontal particle velocity components is given as

$$
\frac{\partial v_x}{\partial t} + \gamma_x v_x = b\frac{\partial p}{\partial x} \tag{17}
$$

Follow the similar derivations we have done for v_z, the final form of time stepping for v_x is given as

$$
\begin{aligned}
v_x[i, j+\frac{1}{2}, it+1] &= \frac{2 - dt\,\gamma_x[j+\frac{1}{2}]}{2 + dt\,\gamma_x[j+\frac{1}{2}]} v_x[i, j+\frac{1}{2}, it] \\
&+ \frac{2b[i, j+\frac{1}{2}]\,dt}{2 + dt\,\gamma_x[j+\frac{1}{2}]} \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(p[i, j+n, it+\frac{1}{2}] - p[i, j-n+1, it+\frac{1}{2}] \right) \right)
\end{aligned} \tag{18}
$$

We can also define diagonal matrix $\mathbf{M}_{vx}\mathbf{B}_{vx}$ and $\mathbf{M}_{vx}\mathbf{B}_p$ and sparse matrix $\mathbf{d}_p\mathbf{d}_x$, all of them are size of $N_z \cdot N_x \times N_z \cdot N_x$. Equation 18 can be simplified as

$$
\mathbf{v}_x^{it+1} = \mathbf{M}_{vx}\mathbf{B}_{vx} \cdot \mathbf{v}_x^{it} + \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x \cdot \mathbf{p}^{it+\frac{1}{2}} \tag{19}
$$

Outside of the PML area, the damping coefficient is 0, the time-stepping function is reduced to

$$
\begin{aligned}
v_x[i, j+\frac{1}{2}, it+1] &= v_x[i, j+\frac{1}{2}, it] \\
&+ b[i, j+\frac{1}{2}]\,dt \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(p[i, j+n, it+\frac{1}{2}] - p[i, j-n+1, it+\frac{1}{2}] \right) \right)
\end{aligned} \tag{20}
$$

This formulation is used for wave field reconstruction backward or forward when the boundary value and the wave field at the last time step is savd. As the size the model is shrinked when the PML layers are trimed, we need to create another diagonal matrix called $\mathbf{R}_{vx}\mathbf{B}p$, whose diagonal elements save the result of $b[j+\frac{1}{2}, j]\,dt$. The size of the diagonal matrix $\mathbf{R}_{vx}\mathbf{B}_p$ is $n_z \cdot n_x \times n_z \cdot n_x$. An new sparse matrix $\mathbf{r}_p\mathbf{d}_x$ is created for approximating the horizontal partial derivative of v_x outside of PML boundary layers. The size of $\mathbf{r}_p\mathbf{d}_x$ is also $n_z \cdot n_x \times n_z \cdot n_x$. With these defined sparse matrix, equation 20 can be written into the vectorized form as

$$
\mathbf{v}_x^{it+1} = \mathbf{v}_x^{it} + \mathbf{R}_{vx}\mathbf{B}_p \cdot \mathbf{r}_p\mathbf{d}_x \cdot \mathbf{p}^{it+\frac{1}{2}} \tag{21}
$$

As before, the forward and adjoint partial derivative operator is given as

$$
\begin{bmatrix}
\frac{\partial p}{\partial x}[1+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[2+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[3+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[4+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[5+\frac{1}{2}]
\end{bmatrix}
= \frac{1}{dx}
\begin{bmatrix}
-a_1 & a_1 & a_2 & & \\
-a_2 & -a_1 & a_1 & a_2 & \\
& -a_2 & -a_1 & a_1 & a_2 \\
& & -a_2 & -a_1 & a_1 \\
& & & -a_2 & -a_1
\end{bmatrix}
\begin{bmatrix}
p[1] \\
p[2] \\
p[3] \\
p[4] \\
p[5]
\end{bmatrix}
\tag{22}
$$

and

$$
\begin{bmatrix}
p[1] \\
p[2] \\
p[3] \\
p[4] \\
p[5]
\end{bmatrix}
= -\frac{1}{dx}
\begin{bmatrix}
a_1 & a_2 & & & \\
-a_1 & a_1 & a_2 & & \\
-a_2 & -a_1 & a_1 & a_2 & \\
& -a_2 & -a_1 & a_1 & a_2 \\
& & -a_2 & -a_1 & a_1
\end{bmatrix}
\begin{bmatrix}
\frac{\partial p}{\partial x}[1+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[2+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[3+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[4+\frac{1}{2}] \\
\frac{\partial p}{\partial x}[5+\frac{1}{2}]
\end{bmatrix}
\tag{23}
$$

$\mathbf{p}_z$ time stepping

The partial derivative equation related p_z is given as

$$
\frac{\partial p_z}{\partial t} + \gamma_z p_z = k \frac{\partial v_z}{\partial z},
\tag{24}
$$

where k is the bulk modulus with $k = \rho v^2$. The partial derivative operator is approximated by finite difference and the center of both side of the equation is on $[i, j, it + 1]$.

$$
\frac{p_z[i,j,it+\frac{3}{2}] - p_z[i,j,it+\frac{1}{2}]}{dt} + \gamma_z[i] \frac{p_z[i,j,it+\frac{3}{2}] + p_z[i,j,it+\frac{1}{2}]}{2}
$$
$$
= k[i,j] \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(v_z[i+n-\frac{1}{2}, j, it+1] - v_z[i-n+\frac{1}{2}, j, it+1] \right) \right)
\tag{25}
$$

After some simple operations, the equation 25 can be simplified as

$$
p_z[i,j,it+\frac{3}{2}] = \frac{2 - dt\,\gamma_z[i]}{2 + dt\,\gamma_z[i]} p_z[i,j,it+\frac{1}{2}]
$$
$$
+ \frac{2k[i,j]\,dt}{2 + dt\,\gamma_z[i]} \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(v_z[i+n-\frac{1}{2}, j, it+1] - v_z[i-n+\frac{1}{2}, j, it+1] \right) \right)
\tag{26}
$$

We create diagonal matrices $\mathbf{M}_{pz}\mathbf{B}_{pz}$ and $\mathbf{M}_{pz}\mathbf{B}_{vz}$ size of $N_z \cdot N_x \times N_z \cdot N_x$, which saves the elements of pr-computed $\frac{2-dt\,\gamma_z[i]}{2+dt\,\gamma_z[i]}$ and $\frac{2k[i,j]\,dt}{2+dt\,\gamma_z[i]}$. We also compute a sparse matrix called $\mathbf{d}_v\mathbf{d}_z$ which facilitate the computation of the vertical partial derivative. With these sparse matrices, the vectorized form of equation 26 can be expressed as

$$
\mathbf{p}_z^{it+\frac{3}{2}} = \mathbf{M}_{pz}\mathbf{B}_{pz} \cdot \mathbf{p}_z^{it+\frac{1}{2}} + \mathbf{M}_{pz}\mathbf{B}_{vz} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it+1}
\tag{27}
$$

Outside of PML absorbing boundary area, the coefficients γ_z is 0, so the equation 26 is reduced to

$$
p_z[i,j,it+\frac{3}{2}] = p_z[i,j,it+\frac{1}{2}]
$$
$$
+ k[i,j]\,dt \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(v_z[i+n-\frac{1}{2}, j, it+1] - v_z[i-n+\frac{1}{2}, j, it+1] \right) \right)
\tag{28}
$$

With respect to wave field reconstruction, there is diagonal matrix called $\mathbf{R}_{pz}\mathbf{B}_{vz}$ save the elements of $k[i,j]\,dt$. Another sparse matrix called $\mathbf{r}_v\mathbf{d}_z$ save coefficients of vertical partial derivative operator. With those expressions, the vectorized form of equation 28 can be written as

$$\mathbf{p}_z^{it+\frac{3}{2}} = \mathbf{p}_z^{it+\frac{1}{2}} + \mathbf{R}_{pz}\mathbf{B}_{vz} \cdot \mathbf{r}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it+1} \tag{29}$$

The vertical partial derivative operator and its adjoint are given as

$$
\begin{bmatrix}
\frac{\partial v_z}{\partial z}[1] \\
\frac{\partial v_z}{\partial z}[2] \\
\frac{\partial v_z}{\partial z}[3] \\
\frac{\partial v_z}{\partial z}[4] \\
\frac{\partial v_z}{\partial z}[5]
\end{bmatrix}
= \frac{1}{dz}
\begin{bmatrix}
a_1 & a_2 & & & \\
-a_1 & a_1 & a_2 & & \\
-a_2 & -a_1 & a_1 & a_2 & \\
 & -a_2 & -a_1 & a_1 & a_2 \\
 & & -a_2 & -a_1 & a_1
\end{bmatrix}
\begin{bmatrix}
v_z[1+\frac{1}{2}] \\
v_z[2+\frac{1}{2}] \\
v_z[3+\frac{1}{2}] \\
v_z[4+\frac{1}{2}] \\
v_z[5+\frac{1}{2}]
\end{bmatrix}
\tag{30}
$$

The adjoint operator is given as

$$
\begin{bmatrix}
v_z[1+\frac{1}{2}] \\
v_z[2+\frac{1}{2}] \\
v_z[3+\frac{1}{2}] \\
v_z[4+\frac{1}{2}] \\
v_z[5+\frac{1}{2}]
\end{bmatrix}
= -\frac{1}{dz}
\begin{bmatrix}
-a_1 & a_1 & a_2 & & \\
-a_2 & -a_1 & a_1 & a_2 & \\
 & -a_2 & -a_1 & a_1 & a_2 \\
 & & -a_2 & -a_1 & a_1 \\
 & & & -a_2 & -a_1
\end{bmatrix}
\begin{bmatrix}
\frac{\partial v_z}{\partial z}[1] \\
\frac{\partial v_z}{\partial z}[2] \\
\frac{\partial v_z}{\partial z}[3] \\
\frac{\partial v_z}{\partial z}[4] \\
\frac{\partial v_z}{\partial z}[5]
\end{bmatrix}
\tag{31}
$$

$\mathbf{p}_x$ time-stepping

$$\frac{\partial p_x}{\partial t} + \gamma_x p_x = k\frac{\partial v_x}{\partial x} + f(t) \tag{32}$$

The finite difference form of this equation is given as

$$p_x[i,j,it+\frac{3}{2}] = \frac{2 - dt\,\gamma_x[j]}{2 + dt\,\gamma_x[j]} p_x[i,j,it+\frac{1}{2}]$$
$$+ \frac{2k[i,j]\,dt}{2 + dt\,\gamma_x[j]} \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(v_x[i,j+n-\frac{1}{2},it+1] - v_x[i,j-n+\frac{1}{2},it+1] \right) \right) + dt \cdot f[k+1]$$

$$\tag{33}$$

we create diagonal matrices $\mathbf{M}_{px}\mathbf{B}_{px}$ and $\mathbf{M}_{px}\mathbf{B}_{vx}$, both are size of $N_z \cdot N_x \times N_z \cdot N_x$, which save the elements of pre-computed $\frac{2-dt\,\gamma_x[j]}{2+dt\,\gamma_x[j]}$ and $\frac{2k[i,j]\,dt}{2+dt\,\gamma_x[j]}$. With the help of these sparse matrices, the vectorized form of equation 33 can be written as

$$\mathbf{p}_x^{it+\frac{3}{2}} = \mathbf{M}_{px}\mathbf{B}_{px} \cdot \mathbf{p}_x^{it+\frac{1}{2}} + \mathbf{M}_{px}\mathbf{B}_{vx} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it+1} + dt \cdot \mathbf{f}^{it+1} \tag{34}$$

Outside of the PML absorbing area, the damping coefficients γ_x is reduced to 0, so the equation 34 can be simplified as

$$p_x[i,j,it+\frac{3}{2}] = p_x[i,j,it+\frac{1}{2}]$$
$$+ k[i,j]\,dt \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(v_x[i,j+n-\frac{1}{2},it+1] - v_x[i,j-n+\frac{1}{2},it+1] \right) \right) + dt \cdot f[k+1]$$

$$\tag{35}$$

We also define sparse matrices $\mathbf{R}_{px}\mathbf{B}_{vx}$ and $\mathbf{r}_v\mathbf{d}_x$, so the vectorized form of equation 35 can be written as

$$\mathbf{p}_x^{it+\frac{3}{2}} = \mathbf{p}_x^{it+\frac{1}{2}} + \mathbf{R}_{px}\mathbf{B}_{vx} \cdot \mathbf{r}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it+1} + dt \cdot \mathbf{f}^{it+1} \tag{36}$$

To facilitate the program, we also give the details of the partial differential operator as

$$\begin{bmatrix} \frac{\partial v_x}{\partial x}[1] \\ \frac{\partial v_x}{\partial x}[2] \\ \frac{\partial v_x}{\partial x}[3] \\ \frac{\partial v_x}{\partial x}[4] \\ \frac{\partial v_x}{\partial x}[5] \end{bmatrix} = \frac{1}{dx} \begin{bmatrix} a_1 & a_2 & & & \\ -a_1 & a_1 & a_2 & & \\ -a_2 & -a_1 & a_1 & a_2 & \\ & -a_2 & -a_1 & a_1 & a_2 \\ & & -a_2 & -a_1 & a_1 \end{bmatrix} \begin{bmatrix} v_x[1+\frac{1}{2}] \\ v_x[2+\frac{1}{2}] \\ v_x[3+\frac{1}{2}] \\ v_x[4+\frac{1}{2}] \\ v_x[5+\frac{1}{2}] \end{bmatrix} \tag{37}$$

and its adjoint operator

$$\begin{bmatrix} v_x[1+\frac{1}{2}] \\ v_x[2+\frac{1}{2}] \\ v_x[3+\frac{1}{2}] \\ v_x[4+\frac{1}{2}] \\ v_x[5+\frac{1}{2}] \end{bmatrix} = -\frac{1}{dx} \begin{bmatrix} -a_1 & a_1 & a_2 & & \\ -a_2 & -a_1 & a_1 & a_2 & \\ & -a_2 & -a_1 & a_1 & a_2 \\ & & -a_2 & -a_1 & a_1 \\ & & & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} \frac{\partial v_x}{\partial x}[1] \\ \frac{\partial v_x}{\partial x}[2] \\ \frac{\partial v_x}{\partial x}[3] \\ \frac{\partial v_x}{\partial x}[4] \\ \frac{\partial v_x}{\partial x}[5] \end{bmatrix} \tag{38}$$

Summarize the discretized wave equations

The discretized wave-equation is given as

$$v_z[i+\frac{1}{2}, j, it+1] = \frac{2 - dt\,\gamma_z[i+\frac{1}{2}]}{(2 + dt\,\gamma_z[i+\frac{1}{2}])} v_z[i+\frac{1}{2}, j, it]$$

$$+ \frac{2b[i+\frac{1}{2}, j]\,dt}{(2 + dt\,\gamma_z[i+\frac{1}{2}])} \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(p[i+n, j, it+\frac{1}{2}] - p[i-n+1, j, it+\frac{1}{2}] \right) \right)$$

$$v_x[i, j+\frac{1}{2}, it+1] = \frac{2 - dt\,\gamma_x[j+\frac{1}{2}]}{2 + dt\,\gamma_x[j+\frac{1}{2}]} v_x[i, j+\frac{1}{2}, it]$$

$$+ \frac{2b[i, j+\frac{1}{2}]\,dt}{2 + dt\,\gamma_x[j+\frac{1}{2}]} \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(p[i, j+n, it+\frac{1}{2}] - p[i, j-n+1, it+\frac{1}{2}] \right) \right)$$

$$p_z[i, j, it+\frac{3}{2}] = \frac{2 - dt\,\gamma_z[i]}{2 + dt\,\gamma_z[i]} p_z[i, j, it+\frac{1}{2}]$$

$$+ \frac{2k[i, j]\,dt}{2 + dt\,\gamma_z[i]} \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(v_z[i+n-\frac{1}{2}, j, it+1] - v_z[i-n+\frac{1}{2}, j, it+1] \right) \right)$$

$$p_x[i, j, it+\frac{3}{2}] = \frac{2 - dt\,\gamma_x[j]}{2 + dt\,\gamma_x[j]} p_x[i, j, it+\frac{1}{2}]$$

$$+ \frac{2k[i, j]\,dt}{2 + dt\,\gamma_x[j]} \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(v_x[i, j+n-\frac{1}{2}, it+1] - v_x[i, j-n+\frac{1}{2}, it+1] \right) \right) + dt \cdot f[k+1] \tag{39}$$

These equation can be written as the sparse matrix-vector format as

$$\begin{aligned}
\mathbf{v}_z^{it+1} &= \mathbf{M}_{vz}\mathbf{B}_{vz} \cdot \mathbf{v}_z^{it} + \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z \cdot \mathbf{p}^{it+\frac{1}{2}} \\
\mathbf{v}_x^{it+1} &= \mathbf{M}_{vx}\mathbf{B}_{vx} \cdot \mathbf{v}_x^{it} + \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x \cdot \mathbf{p}^{it+\frac{1}{2}} \\
\mathbf{p}_z^{it+\frac{3}{2}} &= \mathbf{M}_{pz}\mathbf{B}_{pz} \cdot \mathbf{p}_z^{it+\frac{1}{2}} + \mathbf{M}_{pz}\mathbf{B}_{vz} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it+1} \\
\mathbf{p}_x^{it+\frac{3}{2}} &= \mathbf{M}_{px}\mathbf{B}_{px} \cdot \mathbf{p}_x^{it+\frac{1}{2}} + \mathbf{M}_{px}\mathbf{B}_{vx} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it+1} + dt \cdot \mathbf{f}^{it+1}
\end{aligned} \tag{40}$$

We define the the state variables, which is obtained by vertical concatenate the wave field, at two neighbouring time steps as

$$\mathbf{u}^{it} = \begin{bmatrix} \mathbf{v}_z^{it} \\ \mathbf{v}_x^{it} \\ \mathbf{p}_z^{it+\frac{1}{2}} \\ \mathbf{p}_x^{it+\frac{1}{2}} \end{bmatrix}, \quad \mathbf{u}^{it+1} = \begin{bmatrix} \mathbf{v}_z^{it+1} \\ \mathbf{v}_x^{it+1} \\ \mathbf{p}_z^{it+\frac{3}{2}} \\ \mathbf{p}_x^{it+\frac{3}{2}} \end{bmatrix}, \tag{41}$$

With these expressions, one step forward modelling expressed in equation 40 can be organized into block matrix time block vector form as follow

$$\begin{bmatrix} \mathbf{v}_z^{it+1} \\ \mathbf{v}_x^{it+1} \\ \mathbf{p}_z^{it+\frac{1}{2}} \\ \mathbf{p}_x^{it+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{0} & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_z^{it} \\ \mathbf{v}_x^{it} \\ \mathbf{p}_z^{it+\frac{1}{2}} \\ \mathbf{p}_x^{it+\frac{1}{2}} \end{bmatrix} \tag{42}$$

and

$$\begin{bmatrix} \mathbf{v}_z^{it+1} \\ \mathbf{v}_x^{it+1} \\ \mathbf{p}_z^{it+1+\frac{1}{2}} \\ \mathbf{p}_x^{it+1+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_{pz}\mathbf{B}_{vz} \cdot \mathbf{d}_v\mathbf{d}_z & \mathbf{0} & \mathbf{M}_{pz}\mathbf{B}_{pz} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{vx} \cdot \mathbf{d}_v\mathbf{d}_x & \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{px} \end{bmatrix} \begin{bmatrix} \mathbf{v}_z^{it+1} \\ \mathbf{v}_x^{it+1} \\ \mathbf{p}_z^{it+\frac{1}{2}} \\ \mathbf{p}_x^{it+\frac{1}{2}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ dt \cdot \mathbf{f}^{it+1} \end{bmatrix}. \tag{43}$$

Equation 42 and 43 can be simplified as

$$\mathbf{u}^{it+1} = \mathbf{L}\mathbf{u}^{it} + dt \cdot \mathbf{f}^{it+1}. \tag{44}$$

where

$$\mathbf{L} = \mathbf{L_p}\mathbf{L_v}$$

. The multi-step forward modelling can be expressed as, here I just use four time steps as an example.

$$\begin{bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \mathbf{u}^3 \\ \mathbf{u}^4 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & \mathbf{L} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & \mathbf{L} & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & & & \\ \mathbf{L} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \\ \mathbf{f}^4 \end{bmatrix} \tag{45}$$

Where $\mathbf{f}^{it}$ represented the source-signature at it^{th} time step. $\mathbf{I}$ is a identity matrix size of $4 \cdot N_z \cdot N_x \times 4 \cdot N_z \cdot N_x$. Without the PML boundary condition, all the PML coefficients

$\gamma_z = \gamma_x = 0$, the equation is reduced as

$$
v_z[i+\tfrac{1}{2},j,it+1] = v_z[i+\tfrac{1}{2},j,it] + b[i+\tfrac{1}{2},j]\,dt\left(\frac{1}{dz}\sum_{n=1}^{N}a_n\left(p[i+n,j,it+\tfrac{1}{2}]-p[i-n+1,j,it+\tfrac{1}{2}]\right)\right)
$$

$$
v_x[i,j+\tfrac{1}{2},it+1] = v_x[i,j+\tfrac{1}{2},it] + b[i,j+\tfrac{1}{2}]\,dt\left(\frac{1}{dx}\sum_{n=1}^{N}a_n\left(p[i,j+n,it+\tfrac{1}{2}]-p[i,j-n+1,it+\tfrac{1}{2}]\right)\right)
$$

$$
p_z[i,j,it+\tfrac{3}{2}] = p_z[i,j,it+\tfrac{1}{2}] + k[i,j]\,dt\left(\frac{1}{dz}\sum_{n=1}^{N}a_n\left(v_z[i+n-\tfrac{1}{2},j,it+1]-v_z[i-n+\tfrac{1}{2},j,it+1]\right)\right)
$$

$$
p_x[i,j,it+\tfrac{3}{2}] = p_x[i,j,it+\tfrac{1}{2}] + k[i,j]\,dt\left(\frac{1}{dx}\sum_{n=1}^{N}a_n\left(v_x[i,j+n-\tfrac{1}{2},it+1]-v_x[i,j-n+\tfrac{1}{2},it+1]\right)\right)
$$

$$
+\,dt\cdot f[k+1]
\tag{46}
$$

The vectorized form of above equation can be written as

$$
\begin{aligned}
\mathbf{v}_z^{it+1} &= \mathbf{v}_z^{it} + \mathbf{R}_{vz}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_z\cdot\mathbf{p}^{it+\frac{1}{2}}\\
\mathbf{v}_x^{it+1} &= \mathbf{v}_x^{it} + \mathbf{R}_{vx}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_x\cdot\mathbf{p}^{it+\frac{1}{2}}\\
\mathbf{p}_z^{it+\frac{3}{2}} &= \mathbf{p}_z^{it+\frac{1}{2}} + \mathbf{R}_{pz}\mathbf{B}_{vz}\cdot\mathbf{r}_v\mathbf{d}_z\cdot\mathbf{v}_z^{it+1}\\
\mathbf{p}_x^{it+\frac{3}{2}} &= \mathbf{p}_x^{it+\frac{1}{2}} + \mathbf{R}_{px}\mathbf{B}_{vx}\cdot\mathbf{r}_v\mathbf{d}_x\cdot\mathbf{v}_x^{it+1} + dt\cdot\mathbf{f}^{it+1}
\end{aligned}
\tag{47}
$$

The last two equations in the equation 47 can be combined together to get one equation as the diagonal matrix $\mathbf{R}_{pz}\mathbf{B}_{vz}$ is equal to $\mathbf{R}_{px}\mathbf{B}_{vx}$. So we get the final form as

$$
\begin{aligned}
\mathbf{v}_z^{it+1} &= \mathbf{v}_z^{it} + \mathbf{R}_{vz}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_z\cdot\mathbf{p}^{it+\frac{1}{2}}\\
\mathbf{v}_x^{it+1} &= \mathbf{v}_x^{it} + \mathbf{R}_{vx}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_x\cdot\mathbf{p}^{it+\frac{1}{2}}\\
\mathbf{p}^{it+\frac{3}{2}} &= \mathbf{p}^{it+\frac{1}{2}} + \mathbf{R}_p\mathbf{B}_v\cdot\left(\mathbf{r}_v\mathbf{d}_z\cdot\mathbf{v}_z^{it+1}+\mathbf{r}_v\mathbf{d}_x\cdot\mathbf{v}_x^{it+1}\right) + dt\cdot\mathbf{f}^{it+1}
\end{aligned}
\tag{48}
$$

where $\mathbf{R}_p\mathbf{B}_v = \mathbf{R}_{pz}\mathbf{B}_{vz} = \mathbf{R}_{px}\mathbf{B}_{vx}$ and $\mathbf{p} = \mathbf{p}_z + \mathbf{p}_x$. In a summary, we have created 19 sparse matrices for the acoustic modelling package, they are

$$
\begin{bmatrix}
\mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{M}_{vz}\mathbf{B}_p & \mathbf{d}_p\mathbf{d}_z & \mathbf{R}_{vz}\mathbf{B}_p & \mathbf{r}_p\mathbf{d}_z\\
\mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{M}_{vx}\mathbf{B}_p & \mathbf{d}_p\mathbf{d}_x & \mathbf{R}_{vx}\mathbf{B}_p & \mathbf{r}_p\mathbf{d}_x\\
\mathbf{M}_{pz}\mathbf{B}_{pz} & \mathbf{M}_{pz}\mathbf{B}_{vz} & \mathbf{d}_v\mathbf{d}_z & & \mathbf{r}_v\mathbf{d}_z\\
\mathbf{M}_{px}\mathbf{B}_{px} & \mathbf{M}_{px}\mathbf{B}_{vx} & \mathbf{d}_v\mathbf{d}_x & \mathbf{R}_p\mathbf{B}_v & \mathbf{r}_v\mathbf{d}_x
\end{bmatrix}
\tag{49}
$$

Equation 48 is used for forward reconstruction of the wave field when the boundary value are saved. We also need to reconstruct the wave field backward when we applying the zero-lag cross-correlation imaging conditions to get the gradient or images. Here, we give the backward reconstruction equations

$$
\begin{aligned}
\mathbf{p}^{it+\frac{1}{2}} &= \mathbf{p}^{it+\frac{3}{2}} - \mathbf{R}_p\mathbf{B}_v\cdot\left(\mathbf{r}_v\mathbf{d}_z\cdot\mathbf{v}_z^{it+1}+\mathbf{r}_v\mathbf{d}_x\cdot\mathbf{v}_x^{it+1}\right) - dt\cdot\mathbf{f}^{it+1}\\
\mathbf{v}_x^{it} &= \mathbf{v}_x^{it+1} - \mathbf{R}_{vx}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_x\cdot\mathbf{p}^{it+\frac{1}{2}}\\
\mathbf{v}_z^{it} &= \mathbf{v}_z^{it+1} - \mathbf{R}_{vz}\mathbf{B}_p\cdot\mathbf{r}_p\mathbf{d}_z\cdot\mathbf{p}^{it+\frac{1}{2}}
\end{aligned}
\tag{50}
$$

Gradient derivation

The synthetic data at one(it^{th}) time step $\mathbf{d}^{it} = \mathbf{s}\mathbf{u}^{it}$, where $\mathbf{s}$ is the sampling operator which specified by the acquisition geometry. For FWI, we estimate the model parameters by minimizing the cost function

$$J = \frac{1}{2}||\mathbf{d} - \mathbf{d}_{obs}||_2^2, \tag{51}$$

where $\mathbf{d}$ includes all the times steps, $\mathbf{d} = \begin{bmatrix} \mathbf{d}^{1T} & \mathbf{d}^{2T} & \cdots & \mathbf{d}^{ntT} \end{bmatrix}^T$, $\mathbf{d}_{obs}$ is the observed data. The next part is to derive the gradient of the objective function with respect to one element of the model parameter m_l.

$$\frac{\partial J}{\partial m_l} = (\frac{\partial \mathbf{u}}{\partial m_l})^T \frac{\partial J}{\partial \mathbf{u}} \tag{52}$$

where $\frac{\partial J}{\partial \mathbf{u}} = \mathbf{s}^T(\mathbf{d} - \mathbf{d}_{obs})$. Usually, people name the difference between synthetic data and the observed data as residue $\mathbf{r}$, so we have

$$\frac{\partial J}{\partial \mathbf{u}} = \mathbf{s}^T \mathbf{r} \tag{53}$$

Then we derive the derivative of $\frac{\partial \mathbf{u}}{\partial m_l}$, the identity matrix $\mathbf{I}$ and the source term $\mathbf{f}^{it}$ are independent of model parameter m_l, so their derivative with respect to m_l is zero matrix $\mathbf{0}$.

$$
\frac{\partial \mathbf{u}}{\partial m_l} = \frac{\partial}{\partial m_l}\left(\begin{bmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \mathbf{u}^3 \\ \mathbf{u}^4 \end{bmatrix}\right) = \frac{\partial}{\partial m_l}\left(\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & \mathbf{L} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & \mathbf{L} & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ \mathbf{L} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \\ \mathbf{f}^4 \end{bmatrix}\right)
$$

$$
= \begin{bmatrix} \mathbf{0} & & & \\ & \mathbf{0} & & \\ & & \mathbf{0} & \\ & & \frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & \mathbf{L} & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ \mathbf{L} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \\ \mathbf{f}^4 \end{bmatrix}
$$

$$
+ \begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & \mathbf{L} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{0} & & & \\ & \mathbf{0} & & \\ & \frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{0} & \\ & & & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ \mathbf{L} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \\ \mathbf{f}^4 \end{bmatrix}
$$

$$
+ \begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & \mathbf{L} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & \mathbf{L} & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ \frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{I} & & \\ & & \mathbf{I} & \\ & & & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \mathbf{f}^3 \\ \mathbf{f}^4 \end{bmatrix} \tag{54}
$$

The above expression can be simplified as

$$
\frac{\partial \mathbf{u}}{\partial m_l} =
\begin{bmatrix}
\mathbf{0} & & & \\
& \mathbf{0} & & \\
& & \mathbf{0} & \\
& & \frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}^1 \\
\mathbf{u}^2 \\
\mathbf{u}^3 \\
\mathbf{f}^4
\end{bmatrix}
$$
$$
+
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \\
& & \mathbf{L} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{0} & & & \\
& \mathbf{0} & & \\
& \frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{0} & \\
& & & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}^1 \\
\mathbf{u}^2 \\
\mathbf{f}^3 \\
\mathbf{f}^4
\end{bmatrix} \tag{55}
$$
$$
+
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \\
& & \mathbf{L} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& \mathbf{L} & \mathbf{I} & \\
& & & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{0} & & & \\
\frac{\partial \mathbf{L}}{\partial m_l} & \mathbf{0} & & \\
& & \mathbf{0} & \\
& & & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}^1 \\
\mathbf{f}^2 \\
\mathbf{f}^3 \\
\mathbf{f}^4
\end{bmatrix}
$$

Which can be further simplified as

$$
\frac{\partial \mathbf{u}}{\partial m_l} =
\begin{bmatrix}
\mathbf{0} \\
\mathbf{0} \\
\mathbf{0} \\
\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^3
\end{bmatrix}
$$
$$
+
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \\
& & \mathbf{L} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{0} \\
\mathbf{0} \\
\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^2 \\
\mathbf{0}
\end{bmatrix} \tag{56}
$$
$$
+
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \\
& & \mathbf{L} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& \mathbf{L} & \mathbf{I} & \\
& & & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{0} \\
\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^1 \\
\mathbf{0} \\
\mathbf{0}
\end{bmatrix}
$$

The gradient is given as

$$
\frac{\partial J}{\partial m_l} = \left(\frac{\partial \mathbf{u}}{\partial m_l}\right)^T \frac{\partial J}{\partial \mathbf{u}}
$$
$$
=
\begin{bmatrix}
\mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^3\right)^T
\end{bmatrix}
\begin{bmatrix}
\mathbf{s}^T \mathbf{r}^1 \\
\mathbf{s}^T \mathbf{r}^2 \\
\mathbf{s}^T \mathbf{r}^3 \\
\mathbf{s}^T \mathbf{r}^4
\end{bmatrix}
$$
$$
+
\begin{bmatrix}
\mathbf{0}^T & \mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^2\right)^T & \mathbf{0}^T
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \mathbf{L}^T \\
& & & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{s}^T \mathbf{r}^1 \\
\mathbf{s}^T \mathbf{r}^2 \\
\mathbf{s}^T \mathbf{r}^3 \\
\mathbf{s}^T \mathbf{r}^4
\end{bmatrix} \tag{57}
$$
$$
+
\begin{bmatrix}
\mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l} \mathbf{u}^1\right)^T & \mathbf{0}^T & \mathbf{0}^T
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & \mathbf{L}^T & \\
& & \mathbf{I} & \\
& & & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{I} & & & \\
& \mathbf{I} & & \\
& & \mathbf{I} & \mathbf{L}^T \\
& & & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{s}^T \mathbf{r}^1 \\
\mathbf{s}^T \mathbf{r}^2 \\
\mathbf{s}^T \mathbf{r}^3 \\
\mathbf{s}^T \mathbf{r}^4
\end{bmatrix}
$$

We record the adjoint wave field which generated by backward propagating residues $\mathbf{r}$, as $\mathbf{b}^k$, so the gradient can be simplified as

$$\frac{\partial J}{\partial m_l} = (\frac{\partial \mathbf{u}}{\partial m_l})^T \frac{\partial J}{\partial \mathbf{u}}$$

$$= \begin{bmatrix} \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^3\right)^T \end{bmatrix} \begin{bmatrix} \mathbf{s}^T\mathbf{r}^1 \\ \mathbf{s}^T\mathbf{r}^2 \\ \mathbf{s}^T\mathbf{r}^3 \\ \mathbf{b}^4 \end{bmatrix}$$

$$+ \begin{bmatrix} \mathbf{0}^T & \mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^2\right)^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{s}^T\mathbf{r}^1 \\ \mathbf{s}^T\mathbf{r}^2 \\ \mathbf{b}^3 \\ \mathbf{b}^4 \end{bmatrix}$$

$$+ \begin{bmatrix} \mathbf{0}^T & \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^1\right)^T & \mathbf{0}^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{s}^T\mathbf{r}^1 \\ \mathbf{b}^2 \\ \mathbf{b}^3 \\ \mathbf{b}^4 \end{bmatrix}$$ (58)

$$= \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^3\right)^T \mathbf{b}^4 + \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^2\right)^T \mathbf{b}^3 + \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^1\right)^T \mathbf{b}^2$$

$$= \sum_{it=2}^{nt} \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^{it-1}\right)^T \mathbf{b}^{it}$$

where nt is the total number of time-stepping, $nt = 4$ in this toy example. Let's take a look at $\frac{\partial \mathbf{L}}{\partial m_l}$. To simplify the derivation, we assume the unknown model parameter $\mathbf{m}$ is velocity model in constant density case. As we derived above

$$\mathbf{L} = \mathbf{L}_p \mathbf{L}_v$$ (59)

and

$$\mathbf{L}_p = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_{pz}\mathbf{B}_{vz} \cdot \mathbf{d}_v\mathbf{d}_z & \mathbf{0} & \mathbf{M}_{pz}\mathbf{B}_{pz} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{vx} \cdot \mathbf{d}_v\mathbf{d}_x & \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{px} \end{bmatrix}$$ (60)

and

$$\mathbf{L}_v = \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{0} & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_z \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p\mathbf{d}_x \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}$$ (61)

Only $\mathbf{M}_{pz}\mathbf{B}_{vz}$ and $\mathbf{M}_{px}\mathbf{B}_{vx}$ are depends on the velocity model and both of them are diagonal matrices with elements

$$\mathbf{M}_{pz}\mathbf{B}_{vz}[l, l] = \frac{2k[i, j]\, dt}{2 + dt\, \gamma_z[i]}$$ (62)

and

$$\mathbf{M}_{px}\mathbf{B}_{vx}[l, l] = \frac{2k[i, j]\, dt}{2 + dt\, \gamma_x[j]}$$ (63)

The relationship between linear index l and the two sub-index i, j is

$$l = (j - 1) \times N_z + i \tag{64}$$

If we only consider about the modelling area (outside of PML boundary area), both γ_z and γ_x are equal to 0. so the equation 62 and 63 are reduced to

$$\mathbf{M}_{pz}\mathbf{B}_{vz}[l, l] = \mathbf{M}_{px}\mathbf{B}_{vx}[l, l] = k[i, j] \cdot dt = \rho[i, j] \cdot v[i, j]^2 \cdot dt \tag{65}$$

with the model parameter $m[l] = v[i, j]$. It is easy to get the result $\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l}$ is a square matrix size of $N_z \cdot N_x \times N_z \cdot N_x$, there is only one non-zero element at l^{th} row, l^{th} column, that means

$$\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l}[row_{idx}, col_{idx}] = \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l}[row_{idx}, col_{idx}] = \begin{cases} 2\rho[l] \cdot v[l] \cdot dt, & \text{if } row_{idx} = l \text{ and } col_{idx} = l \\ 0, & \text{otherwise} \end{cases}$$

$$\tag{66}$$

So the partial derivative of the matrix $\mathbf{L}$ with respect to one model paramter m_l is given as

$$\frac{\partial \mathbf{L}}{\partial m_l} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_z & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_x & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{0} & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_z & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_z \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_x & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_x \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{67}$$

So the expression $\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^{it-1}$ is given as

$$\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^{it-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_z & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_x & \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{L}_v \mathbf{u}^{it-1} \tag{68}$$

where $\mathbf{L}_v \mathbf{u}^{it-1}$ is expanded as

$$\mathbf{L}_v \mathbf{u}^{it-1} = \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{0} & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_z & \mathbf{M}_{vz}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_z \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_x & \mathbf{M}_{vx}\mathbf{B}_p \cdot \mathbf{d}_p \mathbf{d}_x \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_z^{it-1} \\ \mathbf{v}_x^{it-1} \\ \mathbf{p}_z^{it-1+\frac{1}{2}} \\ \mathbf{p}_x^{it-1+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_z^{it} \\ \mathbf{v}_x^{it} \\ \mathbf{p}_z^{it-1+\frac{1}{2}} \\ \mathbf{p}_x^{it-1+\frac{1}{2}} \end{bmatrix} \tag{69}$$

With this form, equation 68 can be written as

$$\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}_{k-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_z & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_x & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_z^{it} \\ \mathbf{v}_x^{it} \\ \mathbf{p}_z^{it-1+\frac{1}{2}} \\ \mathbf{p}_x^{it-1+\frac{1}{2}} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_z \cdot \mathbf{v}_z^{it} \\ \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v \mathbf{d}_x \cdot \mathbf{v}_x^{it} \end{bmatrix} \tag{70}$$

The final gradient form expressed in equation 58 can be written as

$$\sum_{it=2}^{nt} \left(\frac{\partial \mathbf{L}}{\partial m_l}\mathbf{u}^{it-1}\right)^T \mathbf{b}^{it} = \sum_{it=2}^{nt} \begin{bmatrix} \mathbf{0}^T & \mathbf{0}^T & \left(\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it}\right)^T & \left(\frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it}\right)^T \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}}_z^{it} \\ \tilde{\mathbf{v}}_x^{it} \\ \tilde{\mathbf{p}}_z^{it+\frac{1}{2}} \\ \tilde{\mathbf{p}}_x^{it+\frac{1}{2}} \end{bmatrix}$$

$$= \sum_{it=2}^{nt} \left(\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it}\right)^T \tilde{\mathbf{p}}_z^{it+\frac{1}{2}} + \left(\frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it}\right)^T \tilde{\mathbf{p}}_x^{it+\frac{1}{2}} \tag{71}$$

Before we simplify the equation 71, we give the expression of $\mathbf{L}^T$

$$\mathbf{L}^T = \mathbf{L}_v^T \mathbf{L}_p^T \tag{72}$$

and $\mathbf{L}_v^T$ is given as

$$\mathbf{L}_v^T = \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz}{}^T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx}{}^T & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p{}^T & \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p{}^T & \mathbf{I} & \mathbf{0} \\ \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p{}^T & \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p{}^T & \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{vz}\mathbf{B}_{vz} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{vx}\mathbf{B}_{vx} & \mathbf{0} & \mathbf{0} \\ \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p & \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p & \mathbf{I} & \mathbf{0} \\ \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p & \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p & \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{73}$$

and the adjoint of $\mathbf{L}_p$ is given as

$$\mathbf{L}_p^T = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{d}_v\mathbf{d}_z{}^T \cdot \mathbf{M}_{pz}\mathbf{B}_{vz}{}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{d}_v\mathbf{d}_x{}^T \cdot \mathbf{M}_{px}\mathbf{B}_{vx}{}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{pz}\mathbf{B}_{pz}{}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{px}{}^T \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{d}_v\mathbf{d}_z{}^T \cdot \mathbf{M}_{pz}\mathbf{B}_{vz} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{d}_v\mathbf{d}_x{}^T \cdot \mathbf{M}_{px}\mathbf{B}_{vx} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_{pz}\mathbf{B}_{pz} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{M}_{px}\mathbf{B}_{px} \end{bmatrix} \tag{74}$$

In the derivation of above two equations, we used a property: the transpose of diagonal matrix is themselves. As we discussed in the equation 28, $\mathbf{M}_{pz}\mathbf{B}_{pz}$ and $\mathbf{M}_{px}\mathbf{B}_{px}$ are identity matrix outside the PML absorbing area, we have

$$\tilde{p}_z^{it-1+\frac{1}{2}} = \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p \cdot \tilde{v}_z^{it-1} + \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p \cdot \tilde{v}_x^{it-1} + \tilde{p}_z^{it+\frac{1}{2}}$$
$$\tilde{p}_z^{it-1+\frac{1}{2}} = \mathbf{d}_p\mathbf{d}_z{}^T \cdot \mathbf{M}_{vz}\mathbf{B}_p \cdot \tilde{v}_z^{it-1} + \mathbf{d}_p\mathbf{d}_x{}^T \cdot \mathbf{M}_{vx}\mathbf{B}_p \cdot \tilde{v}_x^{it-1} + \tilde{p}_z^{it+\frac{1}{2}} \tag{75}$$

So we can get the conclusion that for the adjoint wave field outside the PML absorbing boundary area that

$$\tilde{p}^{it+\frac{1}{2}}/2 = \tilde{p}_z^{it+\frac{1}{2}} = \tilde{p}_x^{it+\frac{1}{2}} \tag{76}$$

With this property, the equation 71 can be further simplified as

$$
\begin{aligned}
\frac{\partial J}{\partial \mathbf{m}_l} &= \sum_{it=2}^{nt} \left(\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} \right)^T \tilde{\mathbf{p}}_z^{it+\frac{1}{2}} + \left(\frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right)^T \tilde{\mathbf{p}}_x^{it+\frac{1}{2}} \\
&= \sum_{it=2}^{nt} \left(\frac{\partial \mathbf{M}_{pz}\mathbf{B}_{vz}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} + \frac{\partial \mathbf{M}_{px}\mathbf{B}_{vx}}{\partial m_l} \cdot \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right)^T \tilde{\mathbf{p}}_x^{it+\frac{1}{2}} \\
&= \sum_{it=2}^{nt} \left(2\rho[l] \cdot v[l] \cdot dt \cdot \left(\mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} \right)[l] + 2\rho[l] \cdot v[l] \cdot dt \cdot \left(\mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right)[l] \right) \cdot \tilde{\mathbf{p}}_x^{it+\frac{1}{2}}[l] \\
&= \sum_{it=2}^{nt} 2\rho[l] \cdot v[l] \cdot dt \cdot \left(\mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} + \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right)[l] \cdot \tilde{\mathbf{p}}_x^{it+\frac{1}{2}}[l]
\end{aligned}
\tag{77}
$$

The gradient of the objective function with respect to all the model parameter is given as

$$
\frac{\partial J}{\partial \mathbf{m}} = \sum_{it=2}^{nt} \mathbf{diagm} \left(2\rho * \mathbf{v} * dt * \left(\mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} + \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right) \right) \tilde{\mathbf{p}}_x^{it+\frac{1}{2}},
\tag{78}
$$

where the symbol $*$ is element-wise multiplication and the operator $\mathbf{diagm}$ representing creating a diagonal matrix from a vector. Summing the last two equations of equation 46, we can get

$$
\begin{aligned}
p[i, j, it + \frac{1}{2}] = {}& p[i, j, it - \frac{1}{2}] + k[i, j]\, dt \left(\frac{1}{dz} \sum_{n=1}^{N} a_n \left(v_z[i + n - \frac{1}{2}, j, it] - v_z[i - n + \frac{1}{2}, j, it] \right) \right) \\
&+ k[i, j]\, dt \left(\frac{1}{dx} \sum_{n=1}^{N} a_n \left(v_x[i, j + n - \frac{1}{2}, it] - v_x[i, j - n + \frac{1}{2}, it] \right) \right) \\
&+ dt \cdot f[it]
\end{aligned}
\tag{79}
$$

From above equations we can get the result

$$
\left(\mathbf{d}_v\mathbf{d}_z \cdot \mathbf{v}_z^{it} + \mathbf{d}_v\mathbf{d}_x \cdot \mathbf{v}_x^{it} \right) = \frac{\mathbf{p}^{it+\frac{1}{2}} - \mathbf{p}^{it-\frac{1}{2}} - dt \cdot \mathbf{f}^{it}}{\mathbf{k} \cdot dt}
\tag{80}
$$

We use the expression in equation 80 to replace the term in equation 77, we can get

$$
\frac{\partial J}{\partial \mathbf{m}} = \sum_{it=2}^{nt} \mathbf{diagm} \left(\frac{2(\mathbf{p}^{it+\frac{1}{2}} - \mathbf{p}^{it-\frac{1}{2}} - dt \cdot \mathbf{f}^{it})}{\mathbf{v}} \right) \tilde{\mathbf{p}}_x^{it+\frac{1}{2}}
\tag{81}
$$

The computations inside the $\mathbf{diagm}$ is element-wise operation.

Born approximation

Researchers commonly use born approximation to derive the gradient for FWI, here we show the result obtained based on matrix formulation is consistent with the gradient derived based

on Born approximation. The first order acoustic wave equation can be expressed as

$$\frac{\partial v_z}{\partial t} = b\frac{\partial p}{\partial z}$$
$$\frac{\partial v_x}{\partial t} = b\frac{\partial p}{\partial x} \tag{82}$$
$$\frac{\partial p}{\partial t} = k\left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x}\right) + f(t)$$

We still assume the density is constant, the scatter wave field is produced by the perturbation of velocity model v, so the new wave field is governed by the following equations:

$$\frac{\partial(v_z + \Delta v_z)}{\partial t} = b\frac{\partial(p + \Delta p)}{\partial z}$$
$$\frac{\partial(v_x + \Delta v_x)}{\partial t} = b\frac{\partial(p + \Delta p)}{\partial x} \tag{83}$$
$$\frac{\partial(p + \Delta p)}{\partial t} = \rho(v + \Delta v)^2\left(\frac{\partial(v_z + \Delta v_z)}{\partial z} + \frac{\partial(v_x + \Delta v_x)}{\partial x}\right) + f(t)$$

Subtracting equation 83 by equation 82 and ignore the higher order terms, we get the scattered wave field equations as

$$\frac{\partial \Delta v_z}{\partial t} = b\frac{\partial}{\partial z}\left(\Delta p_z + \Delta p_x\right)$$
$$\frac{\partial \Delta v_x}{\partial t} = b\frac{\partial}{\partial x}\left(\Delta p_z + \Delta p_x\right) \tag{84}$$
$$\frac{\partial \Delta p}{\partial t} = k\left(\frac{\partial \Delta v_z}{\partial z} + \frac{\partial \Delta v_z}{\partial z}\right) + 2\rho v \Delta v\left(\frac{\partial v_z}{\partial z} + \frac{\partial v_z}{\partial z}\right)$$

From the last equation of equations 82, we know

$$\frac{\partial v_z}{\partial z} + \frac{\partial v_z}{\partial z} = \frac{\frac{\partial p}{\partial t} - f(t)}{k} \tag{85}$$

We can use this equation to replace the terms in equation 83, we can get

$$\frac{\partial \Delta v_z}{\partial t} = b\frac{\partial}{\partial z}\left(\Delta p_z + \Delta p_x\right)$$
$$\frac{\partial \Delta v_x}{\partial t} = b\frac{\partial}{\partial x}\left(\Delta p_z + \Delta p_x\right) \tag{86}$$
$$\frac{\partial \Delta p}{\partial t} = k\left(\frac{\partial \Delta v_z}{\partial z} + \frac{\partial \Delta v_z}{\partial z}\right) + \frac{2\left(\frac{\partial p}{\partial t} - f(t)\right)}{v}\Delta v$$

We usually call the term in front of Δv as the source-side wave filed, which is

$$\frac{2\left(\frac{\partial p}{\partial t} - f(t)\right)}{v} \tag{87}$$

The gradient of the velocity model is obtained by zero-lag cross-correlation between source-side wave field and the adjoint wave field. Which is consistent with the discretized form give in equation 81 as

$$\mathbf{diagm}\left(\frac{2(\mathbf{p}^{it+\frac{1}{2}} - \mathbf{p}^{it-\frac{1}{2}} - dt \cdot \mathbf{f}^{it})}{\mathbf{v}}\right) \tag{88}$$

The only difference between above two equations is dt, which come from the discretization of equation 84, and equation 88 can be obtained by multiplying equation 85 by dt.

Gauss-newton method

Full-newton method

FREQUENCY DOMAIN

We start with the splitted first-order acoustic wave-equation

$$
\begin{aligned}
\frac{\partial v_z}{\partial t} + \gamma_z v_z &= b\frac{\partial(p_z + p_x)}{\partial z} \\
\frac{\partial v_x}{\partial t} + \gamma_x v_x &= b\frac{\partial(p_z + p_x)}{\partial x} \\
\frac{\partial p_z}{\partial t} + \gamma_z p_z &= k\frac{\partial v_z}{\partial z} \\
\frac{\partial p_x}{\partial t} + \gamma_x p_x &= k\frac{\partial v_x}{\partial x} + f(t)
\end{aligned}
\tag{89}
$$

where $\gamma(x) = c_{pml}cos(\frac{\pi}{2}\frac{x}{L})$ where x is the local coordinate in the PML layer, the origin is located at the outer edges of the model, L is the width of PML layer. the scalar c_{pml} is a scalar defined by trial and error. After disrectization of the PML coefficients, both γ_z and γ_x are a 1D vector.

Perform Fourier transform along the time axis to equation 89, we can get the wave-equations in frequency domain

$$
\begin{aligned}
iwv_z + \gamma_z v_z &= b\frac{\partial(p_z + p_x)}{\partial z} \\
iwv_x + \gamma_x v_x &= b\frac{\partial(p_z + p_x)}{\partial x} \\
iwp_z + \gamma_z p_z &= k\frac{\partial v_z}{\partial z} \\
iwp_x + \gamma_x p_x &= k\frac{\partial v_x}{\partial x} + f(w)
\end{aligned}
\tag{90}
$$

where i is the imagary unit and w is the radian frequency $w = 2\pi f$, where f is frequency. Combine the two terms on the left side, we can get

$$
\begin{aligned}
iw(1 + \frac{\gamma_z}{iw})v_z &= b\frac{\partial(p_z + P_x)}{\partial z} \\
iw(1 + \frac{\gamma_x}{iw})v_x &= b\frac{\partial(p_z + p_x)}{\partial x} \\
iw(1 + \frac{\gamma_z}{iw})p_z &= k\frac{\partial v_z}{\partial z} \\
iw(1 + \frac{\gamma_x}{iw})p_x &= k\frac{\partial v_x}{\partial x} + f(w)
\end{aligned}
\tag{91}
$$

We define new variables $\epsilon_z = 1 + \frac{\gamma_z}{iw} = 1 - \frac{i\gamma_z}{w}$ and $\epsilon_x = 1 + \frac{\gamma_x}{iw} = 1 - \frac{i\gamma_x}{w}$. As we have discussed for γ_z and γ_x, ϵ_z and ϵ_x are also two 1D vectors. We divided equation 91 by ϵ_z

or ϵ_x, we can get

$$iwv_z = \frac{b}{\epsilon_z}\frac{\partial p}{\partial z}$$

$$iwv_x = \frac{b}{\epsilon_x}\frac{\partial p}{\partial x}$$

$$\frac{iw}{k}p_z = \frac{1}{\epsilon_z}\frac{\partial v_z}{\partial z} \tag{92}$$

$$\frac{iw}{k}p_x = \frac{1}{\epsilon_x}\frac{\partial v_x}{\partial x} + \frac{f(w)}{k\epsilon_x}$$

Note: the main purpose of moving ϵ_z and ϵ_x to the right side is that we can sum the last two equations to recover the pressure componnent p.

Discretization on stagger grid

We use stagger grid to discretize the wave-equation. we place the pressure component p_z and p_x, bulk modulus k, buoyancy b on integer grid, partical velocity v_z and v_x on half grid along vertical and horizontal direction, respectively. So the discretized field is given as $v_z[i+\frac{1}{2},j]$, $v_x[i,j+\frac{1}{2}]$, $p_z[i,j]$, $p_x[i,j]$, $b[i,j]$, $k[i,j]$. With this discretizatino, assume the grid size is h in both vertical and horizontal directions. The partial derivative is approximated by second-order finite-difference method. so we can get

$$\frac{\partial v_z}{\partial z}[i,j] = \frac{v_z[i+\frac{1}{2},j]-v_z[i-\frac{1}{2},j]}{h}$$

$$\frac{\partial v_x}{\partial x}[i,j] = \frac{v_x[i,j+\frac{1}{2}]-v_z[i,j-\frac{1}{2}]}{h} \tag{93}$$

From the first two equations of equation 92 and moving the iw term from left side of Equation 92 to the right hand side, we can get

$$v_z[i+\frac{1}{2},j] = \frac{b[i+\frac{1}{2},j]}{iw\,\epsilon_z[i+\frac{1}{2}]}\frac{1}{h}(p[i+1,j]-p[i,j])$$

$$v_z[i-\frac{1}{2},j] = \frac{b[i-\frac{1}{2},j]}{iw\,\epsilon_z[i-\frac{1}{2}]}\frac{1}{h}(p[i,j]-p[i-1,j])$$

$$v_x[i,j+\frac{1}{2}] = \frac{b[i,j+\frac{1}{2}]}{iw\,\epsilon_x[j+\frac{1}{2}]}\frac{1}{h}(p[i,j+1]-p[i,j]) \tag{94}$$

$$v_x[i,j-\frac{1}{2}] = \frac{b[i,j-\frac{1}{2}]}{iw\,\epsilon_x[j-\frac{1}{2}]}\frac{1}{h}(p[i,j]-p[i,j-1])$$

Using the expressions in Equation 94 to replace the terms in equation 93, we can get a new system as

$$\frac{\partial v_z}{\partial z}[i,j] = \frac{b[i+\frac{1}{2},j]}{iw\,\epsilon_z[i+\frac{1}{2}]}\frac{1}{h^2}(p[i+1,j]-p[i,j]) - \frac{b[i-\frac{1}{2},j]}{iw\,\epsilon_z[i-\frac{1}{2}]}\frac{1}{h^2}(p[i,j]-p[i-1,j])$$

$$\frac{\partial v_x}{\partial x}[i,j] = \frac{b[i,j+\frac{1}{2}]}{iw\,\epsilon_x[j+\frac{1}{2}]}\frac{1}{h^2}(p[i,j+1]-p[i,j]) - \frac{b[i,j-\frac{1}{2}]}{iw\,\epsilon_x[j-\frac{1}{2}]}\frac{1}{h^2}(p[i,j]-p[i,j-1]) \tag{95}$$

The last two equations of Equation 92 can be discretized as

$$
\frac{iw}{k[i,j]} p_z[i,j] = \frac{1}{\epsilon_z[i]} \frac{\partial v_z}{\partial z}[i,j]
$$
$$
\frac{iw}{k[i,j]} p_x[i,j] = \frac{1}{\epsilon_x[j]} \frac{\partial v_x}{\partial x}[i,j] + \frac{f(w)}{k[i,j]\epsilon_x[j]}
$$

(96)

We know the source is always inject outside of the PML area, so $\gamma_x = 0$ which leads to $\epsilon_x = 1$, so equations 96 can be simplified as

$$
\frac{iw}{k[i,j]} p_z[i,j] = \frac{1}{\epsilon_z[i]} \frac{\partial v_z}{\partial z}[i,j]
$$
$$
\frac{iw}{k[i,j]} p_x[i,j] = \frac{1}{\epsilon_x[j]} \frac{\partial v_x}{\partial x}[i,j] + \frac{f(w)}{k[i,j]}
$$

(97)

Using the expressions in equations 95 to replace the terms in equation 97, then we can get

$$
\frac{iw}{k[i,j]} p_z[i,j] = \frac{1}{iw\,\epsilon_z[i]h^2} \left(\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} (p[i+1,j] - p[i,j]) - \frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]} (p[i,j] - p[i-1,j]) \right)
$$
$$
\frac{iw}{k[i,j]} p_x[i,j] = \frac{1}{iw\,\epsilon_x[j]h^2} \left(\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} (p[i,j+1] - p[i,j]) - \frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]} (p[i,j] - p[i,j-1]) \right) + \frac{f(w)}{k[i,j]}
$$

(98)

Summing above two equations and multiply iw on both side of equations 98, we get

$$
\frac{-w^2}{k[i,j]} p[i,j] = \frac{1}{\epsilon_z[i]h^2} \left(\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} (p[i+1,j] - p[i,j]) - \frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]} (p[i,j] - p[i-1,j]) \right)
$$
$$
+ \frac{1}{\epsilon_x[j]h^2} \left(\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} (p[i,j+1] - p[i,j]) - \frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]} (p[i,j] - p[i,j-1]) \right)
$$
$$
+ \frac{iw\,f(w)}{k[i,j]}
$$

(99)

To facilitate building the Helmhotz operator as a sparse matrix, we can move the pressure component to the right size of the equation and the source term to the left side of the equation, we get

$$
-\frac{1}{\epsilon_z[i]h^2} \left(\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} (p[i+1,j] - p[i,j]) - \frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]} (p[i,j] - p[i-1,j]) \right)
$$
$$
-\frac{1}{\epsilon_x[j]h^2} \left(\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} (p[i,j+1] - p[i,j]) - \frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]} (p[i,j] - p[i,j-1]) \right)
$$
$$
-\frac{w^2}{k[i,j]} p[i,j] = \frac{iw\,f(w)}{k[i,j]}
$$

(100)

Through some simple organization, we get the coefficients in front of each term

$$p[i,j] \rightarrow -\frac{w^2}{k[i,j]} + \frac{1}{\epsilon_z[i]h^2}\left(\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]}\right) + \frac{1}{\epsilon_x[j]h^2}\left(\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} + \frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]}\right)$$

$$p[i+1,j] \rightarrow -\frac{1}{\epsilon_z[i]h^2}\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]}$$

$$p[i-1,j] \rightarrow -\frac{1}{\epsilon_z[i]h^2}\frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]}$$

$$p[i,j+1] \rightarrow -\frac{1}{\epsilon_x[j]h^2}\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]}$$

$$p[i,j-1] \rightarrow -\frac{1}{\epsilon_x[j]h^2}\frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]}$$

$$(101)$$

Rotated stagger grid

This method still put the pressure components on the grid points. However, the new coordinate system is rotated 45 degrees. The distribution of wave field and model parameters are displayed in figure 2. The acoustic wave equation in the rotated coordinate system is given as

$$iwv_z = \frac{\sqrt{2}}{2}\frac{b}{\epsilon_z}\left(\frac{\partial p}{\partial z'} - \frac{\partial p}{\partial x'}\right)$$

$$iwv_x = \frac{\sqrt{2}}{2}\frac{b}{\epsilon_x}\left(\frac{\partial p}{\partial z'} + \frac{\partial p}{\partial x'}\right)$$

$$\frac{iw}{k}p_z = \frac{\sqrt{2}}{2\epsilon_z}\left(\frac{\partial v_z}{\partial z'} - \frac{\partial v_z}{\partial x'}\right)$$

$$\frac{iw}{k}p_x = \frac{\sqrt{2}}{2\epsilon_x}\left(\frac{\partial v_x}{\partial z'} + \frac{\partial v_x}{\partial x'}\right) + \frac{f(w)}{k}$$

$$(102)$$

We first approximate the spatial partial derivative of partical velocity by finite difference, so we have

$$\frac{\partial v_z}{\partial z'}[i,j] = \frac{v_z[i+\frac{1}{2},j+\frac{1}{2}] - v_z[i-\frac{1}{2},j-\frac{1}{2}]}{\sqrt{2}h}$$

$$\frac{\partial v_z}{\partial x'}[i,j] = \frac{v_z[i-\frac{1}{2},j+\frac{1}{2}] - v_z[i+\frac{1}{2},j-\frac{1}{2}]}{\sqrt{2}h}$$

$$\frac{\partial v_x}{\partial z'}[i,j] = \frac{v_x[i+\frac{1}{2},j+\frac{1}{2}] - v_x[i-\frac{1}{2},j-\frac{1}{2}]}{\sqrt{2}h}$$

$$\frac{\partial v_x}{\partial x'}[i,j] = \frac{v_x[i-\frac{1}{2},j+\frac{1}{2}] - v_x[i+\frac{1}{2},j-\frac{1}{2}]}{\sqrt{2}h}$$

$$(103)$$

From the first equation of the system 88:

$$iwv_z = \frac{\sqrt{2}}{2}\frac{b}{\epsilon_z}\left(\frac{\partial p}{\partial z'} - \frac{\partial p}{\partial x'}\right)$$
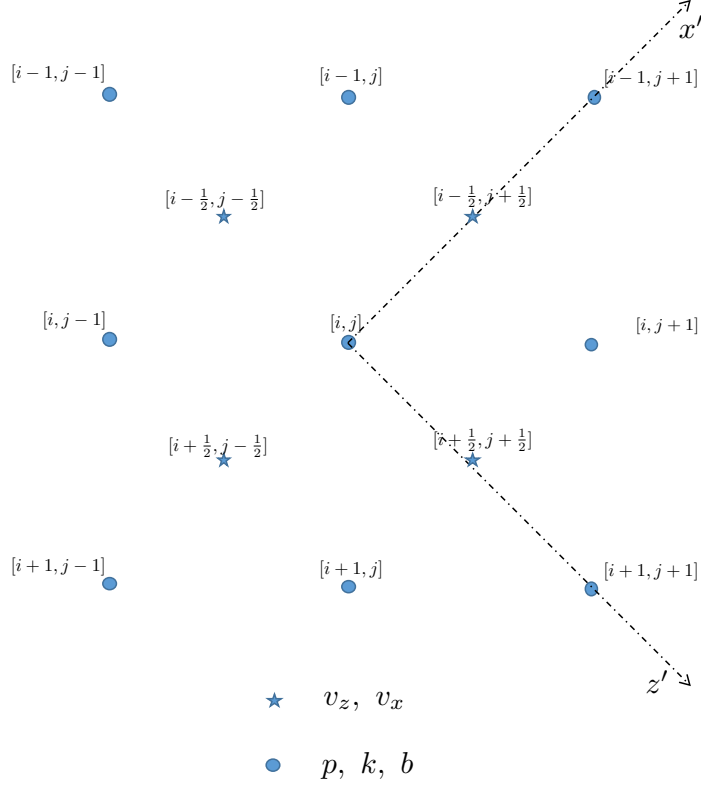
$$(104)$$

Figure 2: Rotated stagger grid

we represent particle velocity component v_z by pressure component p as

$$v_z[i - \frac{1}{2}, j - \frac{1}{2}] = \frac{b[i - \frac{1}{2}, j - \frac{1}{2}]}{2h\,iw\epsilon_z[i - \frac{1}{2}]} \left(p[i,j] - p[i-1,j-1] + p[i,j-1] - p[i-1,j]\right)$$

$$v_z[i + \frac{1}{2}, j - \frac{1}{2}] = \frac{b[i + \frac{1}{2}, j - \frac{1}{2}]}{2h\,iw\epsilon_z[i + \frac{1}{2}]} \left(p[i+1,j] - p[i,j-1] + p[i+1,j-1] - p[i,j]\right)$$

$$v_z[i - \frac{1}{2}, j + \frac{1}{2}] = \frac{b[i - \frac{1}{2}, j + \frac{1}{2}]}{2h\,iw\epsilon_z[i - \frac{1}{2}]} \left(p[i,j+1] - p[i-1,j] + p[i,j] - p[i-1,j+1]\right)$$

$$v_z[i + \frac{1}{2}, j + \frac{1}{2}] = \frac{b[i + \frac{1}{2}, j + \frac{1}{2}]}{2h\,iw\epsilon_z[i + \frac{1}{2}]} \left(p[i+1,j+1] - p[i,j] + p[i+1,j] - p[i,j+1]\right)$$

(105)

From the second equation of system 88:

$$iwv_x = \frac{\sqrt{2}}{2} \frac{b}{\epsilon_x} \left(\frac{\partial p}{\partial z'} + \frac{\partial p}{\partial x'}\right)$$

(106)

By taking advantage of this equation, we can also represent v_x via pressure component p

as:

$$v_x[i - \frac{1}{2}, j - \frac{1}{2}] = \frac{b[i - \frac{1}{2}, j - \frac{1}{2}]}{2h\,iw\epsilon_x[j - \frac{1}{2}]} \left(p[i, j] - p[i - 1, j - 1] + p[i - 1, j] - p[i, j - 1]\right)$$

$$v_x[i + \frac{1}{2}, j - \frac{1}{2}] = \frac{b[i + \frac{1}{2}, j - \frac{1}{2}]}{2h\,iw\epsilon_x[j - \frac{1}{2}]} \left(p[i + 1, j] - p[i, j - 1] + p[i, j] - p[i + 1, j - 1]\right)$$

$$\tag{107}$$

$$v_x[i - \frac{1}{2}, j + \frac{1}{2}] = \frac{b[i - \frac{1}{2}, j + \frac{1}{2}]}{2h\,iw\epsilon_x[j + \frac{1}{2}]} \left(p[i, j + 1] - p[i - 1, j] + p[i - 1, j + 1] - p[i, j]\right)$$

$$v_x[i + \frac{1}{2}, j + \frac{1}{2}] = \frac{b[i + \frac{1}{2}, j + \frac{1}{2}]}{2h\,iw\epsilon_x[j + \frac{1}{2}]} \left(p[i + 1, j + 1] - p[i, j] + p[i, j + 1] - p[i + 1, j]\right)$$

Summing the last two equations of the equation 102, we can get

$$\frac{iw}{k}p = \frac{\sqrt{2}}{2\epsilon_z}\left(\frac{\partial v_z}{\partial z'} - \frac{\partial v_z}{\partial x'}\right) + \frac{\sqrt{2}}{2\epsilon_x}\left(\frac{\partial v_x}{\partial z'} + \frac{\partial v_x}{\partial x'}\right) + \frac{f(w)}{k} \tag{108}$$

Using the expressions in equation 103 to replace the terms in equation **??**, we can get

$$\frac{iw}{k[i, j]}p[i, j] = \frac{1}{2h\epsilon_z[i]}\left(v_z[i + \frac{1}{2}, j + \frac{1}{2}] - v_z[i - \frac{1}{2}, j - \frac{1}{2}] + v_z[i + \frac{1}{2}, j - \frac{1}{2}] - v_z[i - \frac{1}{2}, j + \frac{1}{2}]\right)$$

$$+ \frac{1}{2h\epsilon_x[j]}\left(v_x[i + \frac{1}{2}, j + \frac{1}{2}] - v_x[i - \frac{1}{2}, j - \frac{1}{2}] + v_x[i - \frac{1}{2}, j + \frac{1}{2}] - v_x[i + \frac{1}{2}, j - \frac{1}{2}]\right) + \frac{f(w)}{k[i, j]}$$

$$\tag{109}$$

The next step is to use the expression in equations (106, 107) to replace the particle velocity terms in the equation 109. We also take out the terms $2h$ and iw, then we can get a equation only with pressure component.

$$\frac{-w^2}{k[i, j]}p[i, j] = \frac{b[i + \frac{1}{2}, j + \frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i + \frac{1}{2}]} \left(p[i + 1, j + 1] - p[i, j] + p[i + 1, j] - p[i, j + 1]\right)$$

$$- \frac{b[i - \frac{1}{2}, j - \frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i - \frac{1}{2}]} \left(p[i, j] - p[i - 1, j - 1] + p[i, j - 1] - p[i - 1, j]\right)$$

$$+ \frac{b[i + \frac{1}{2}, j - \frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i + \frac{1}{2}]} \left(p[i + 1, j] - p[i, j - 1] + p[i + 1, j - 1] - p[i, j]\right)$$

$$- \frac{b[i - \frac{1}{2}, j + \frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i - \frac{1}{2}]} \left(p[i, j + 1] - p[i - 1, j] + p[i, j] - p[i - 1, j + 1]\right)$$

$$+ \frac{b[i + \frac{1}{2}, j + \frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j + \frac{1}{2}]} \left(p[i + 1, j + 1] - p[i, j] + p[i, j + 1] - p[i + 1, j]\right) \tag{110}$$

$$- \frac{b[i - \frac{1}{2}, j - \frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j - \frac{1}{2}]} \left(p[i, j] - p[i - 1, j - 1] + p[i - 1, j] - p[i, j - 1]\right)$$

$$+ \frac{b[i - \frac{1}{2}, j + \frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j + \frac{1}{2}]} \left(p[i, j + 1] - p[i - 1, j] + p[i - 1, j + 1] - p[i, j]\right)$$

$$- \frac{b[i + \frac{1}{2}, j - \frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j - \frac{1}{2}]} \left(p[i + 1, j] - p[i, j - 1] + p[i, j] - p[i + 1, j - 1]\right)$$

$$+ \frac{iw\,f(w)}{k[i, j]}$$

To facilitate coding, we group the common terms in equation 110 together, we get a new equations shown as bellow

$$
\begin{aligned}
(\frac{-w^2}{k[i,j]} &+ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} \\
&+ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]})p[i,j] \\
&- (\frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]})p[i+1,j+1] \\
&- (\frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]})p[i-1,j-1] \\
&- (\frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]})p[i+1,j-1] \\
&- (\frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]})p[i-1,j+1] \\
&- (\frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]})p[i+1,j] \\
&- (\frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]})p[i,j+1] \\
&- (\frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]})p[i,j-1] \\
&- (\frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]})p[i-1,j] \\
&= \frac{iw\,f(w)}{k[i,j]}
\end{aligned}
\tag{111}
$$

Mixed grid

By combining the finte-difference grid based on Cartesian and the rotated coordinate system, we can get the mixed-grid finite-difference stencil. This process includes two steps:

Mass averaging

$$
\begin{aligned}
\frac{w^2}{k[i,j]}p[i,j] = w^2(&c\cdot\frac{p[i,j]}{k[i,j]} \\
&+d(\frac{p[i+1,j]}{k[i+1,j]} + \frac{p[i-1,j]}{k[i-1,j]} + \frac{p[i,j+1]}{k[i,j+1]} + \frac{p[i,j-1]}{k[i,j-1]}) \\
&+\frac{1-c-4d}{4}(\frac{p[i+1,j+1]}{k[i+1,j+1]} + \frac{p[i-1,j+1]}{k[i-1,j+1]} + \frac{p[i+1,j-1]}{k[i+1,j-1]} + \frac{p[i-1,j-1]}{k[i-1,j-1]}))
\end{aligned}
\tag{112}
$$

Weighted averaging of the two stencil

The weighting between the two grid is determined by parameter α

$$-\frac{w^2}{k}p - a \cdot \left(\frac{\partial}{\partial z}(b\frac{\partial}{\partial z}) + \frac{\partial}{\partial x}(b\frac{\partial}{\partial x}) \right) p - (1-a) \cdot \left(\frac{\partial}{\partial z'}(b\frac{\partial}{\partial z'}) + \frac{\partial}{\partial x'}(b\frac{\partial}{\partial x'}) \right) p = \frac{iw\,f(w)}{k} \quad (113)$$

where $c = 0.6287326$, $d = 0.092816675$, $\alpha = 0.5617365$. so the coefficient $\frac{1-c-4d}{4} = 1.75e-7$.
After some simple re-organzation of the terms, we give the coefficients for the 9 grid points

$$p[i,j] \rightarrow -c\frac{w^2}{k[i,j]} + \alpha \left\{ \frac{1}{\epsilon_z[i]h^2} \left(\frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]} \right) + \frac{1}{\epsilon_x[j]h^2} \left(\frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} + \frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]} \right) \right\}$$

$$+(1-\alpha) \left\{ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} \right\}$$

$$+(1-\alpha) \left\{ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} \right\}$$

$$(114)$$

$$p[i-1,j-1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i-1,j-1]} - (1-\alpha) \left\{ \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} \right\}$$

$$(115)$$

$$p[i,j-1] \rightarrow -d\frac{w^2}{k[i,j-1]} - \alpha\frac{1}{\epsilon_x[j]h^2}\frac{b[i,j-\frac{1}{2}]}{\epsilon_x[j-\frac{1}{2}]}$$

$$-(1-\alpha) \left\{ \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} \right\}$$

$$(116)$$

$$p[i+1,j-1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i+1,j-1]} - (1-\alpha) \left\{ \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} \right\}$$

$$(117)$$

$$p[i-1,j] \rightarrow -d\frac{w^2}{k[i-1,j]} - \alpha \left\{ \frac{1}{\epsilon_z[i]h^2}\frac{b[i-\frac{1}{2},j]}{\epsilon_z[i-\frac{1}{2}]} \right\}$$

$$-(1-\alpha) \left\{ \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} \right\}$$

$$(118)$$

$$p[i+1,j] \rightarrow -d\frac{w^2}{k[i+1,j]} - \alpha \left\{ \frac{1}{\epsilon_z[i]h^2} \frac{b[i+\frac{1}{2},j]}{\epsilon_z[i+\frac{1}{2}]} \right\}$$

$$-(1-\alpha)\left\{ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j-\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j-\frac{1}{2}]} \right\}$$

$$(119)$$

$$p[i-1,j+1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i-1,j+1]} - (1-\alpha)\left\{ \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} + \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} \right\}$$

$$(120)$$

$$p[i,j+1] \rightarrow -d\frac{w^2}{k[i,j+1]} - \alpha \left\{ \frac{1}{\epsilon_x[j]h^2} \frac{b[i,j+\frac{1}{2}]}{\epsilon_x[j+\frac{1}{2}]} \right\}$$

$$-(1-\alpha)\left\{ \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} - \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} - \frac{b[i-\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i-\frac{1}{2}]} \right\}$$

$$(121)$$

$$p[i+1,j+1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i+1,j+1]} - (1-\alpha)\left\{ \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_z[i]\epsilon_z[i+\frac{1}{2}]} + \frac{b[i+\frac{1}{2},j+\frac{1}{2}]}{4h^2\epsilon_x[j]\epsilon_x[j+\frac{1}{2}]} \right\}$$

$$(122)$$

Outside of PML area, $\epsilon_z = \epsilon_x = 1$, the above coefficients can be simplified as

$$p[i,j] \rightarrow -c\frac{w^2}{k[i,j]} + \frac{a}{h^2}(b[i+\frac{1}{2},j] + b[i-\frac{1}{2},j] + b[i,j+\frac{1}{2}] + b[i,j-\frac{1}{2}])$$

$$+ \frac{1-a}{2h^2}(b[i+\frac{1}{2},j+\frac{1}{2}] + b[i-\frac{1}{2},j-\frac{1}{2}] + b[i-\frac{1}{2},j+\frac{1}{2}] + b[i+\frac{1}{2},j-\frac{1}{2}])$$

$$p[i-1,j-1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i-1,j-1]} - \frac{1-a}{2h^2}b[i-\frac{1}{2},j-\frac{1}{2}]$$

$$p[i,j-1] \rightarrow -d\frac{w^2}{k[i,j-1]} - \frac{\alpha}{h^2}b[i,j-\frac{1}{2}]$$

$$p[i+1,j-1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i+1,j-1]} - \frac{1-a}{2h^2}b[i+\frac{1}{2},j-\frac{1}{2}]$$

$$p[i-1,j] \rightarrow -d\frac{w^2}{k[i-1,j]} - \frac{\alpha}{h^2}b[i-\frac{1}{2},j]$$

$$p[i+1,j] \rightarrow -d\frac{w^2}{k[i+1,j]} - \frac{\alpha}{h^2}b[i+\frac{1}{2},j]$$

$$p[i-1,j+1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i-1,j+1]} - \frac{1-a}{2h^2}b[i-\frac{1}{2},j+\frac{1}{2}]$$

$$p[i,j+1] \rightarrow -d\frac{w^2}{k[i,j+1]} - \frac{\alpha}{h^2}b[i,j+\frac{1}{2}]$$

$$p[i+1,j+1] \rightarrow -\frac{1-c-4d}{4}\frac{w^2}{k[i+1,j+1]} - \frac{1-a}{2h^2}b[i+\frac{1}{2},j+\frac{1}{2}]$$

$$(123)$$

The expression in equation 123 can be used to derive the gradient of FWI for model parameters. The source wavelet is add as

$$\frac{iwf(w)}{k[i,j]},\qquad (124)$$

where $[i,j]$ indicate the location of source point.

Gradient for FWI in frequency domain

Based on above formulations, we can create the Helmholtz operator as an sparse matrix $\mathbf{H}$, so the forward modelling process can be expressed as

$$\mathbf{H(m)\,u(m)} = \mathbf{f} \qquad (125)$$

where $\mathbf{H}$ is the Helmholtz operator and $\mathbf{m}$ is model parameter, $\mathbf{u}$ is the wave field at one frequency component and $\mathbf{f}$ is the source signature in the frequency domain. Follow the notations we used for time-domain finite difference method, the size of $\mathbf{H}$ is $N_z \cdot N_x \times N_z \cdot N_x$. The wave field at one frequency slice $\mathbf{u}$ is a vector length of $N_z \cdot N_x$, so as to the source signature $\mathbf{f}$.

To simplify the derivations of gradient, we temporarily just consider one frequency slice. The synthetic data is obtained by sampling the wave field at the receiver's location, the sampling operator is indicated by operator $\mathbf{s}$. So the synthetic data is represented as

$$\mathbf{d} = \mathbf{su} \qquad (126)$$

In FWI, the model parameter is updated by minimizing the difference between the synthetic and observed data, here we just use the $l2$ norm as the metric to evaluate the difference, so the objective function is given as

$$J = \frac{1}{2}||\mathbf{d} - \mathbf{d}_{obs}||_2^2 = \frac{1}{2}(\mathbf{d} - \mathbf{d}_{obs})^T(\mathbf{d} - \mathbf{d}_{obs})^* \tag{127}$$

Where $\mathbf{d}_{obs}$ is the observed data, the symbol ^{T} and * represent matrix transpose and complex conjugate. We assume the density is constant, so the model parameter is just the velocity model.

We derive the derivative of the cost with respect to one model parameter m_l, l is the linear index of this model parameter. The derivative is given as

$$\frac{\partial J}{\partial m_l} = \left(\frac{\partial \mathbf{u}}{\partial m_l}\right)^T \frac{\partial J}{\partial \mathbf{u}} + \left(\frac{\partial \mathbf{u}^*}{\partial m_l}\right)^T \frac{\partial J}{\partial \mathbf{u}^*} \tag{128}$$

Base on the equation 125, 126, using the property of complex conjugate, we can get

$$\mathbf{H}^*(\mathbf{m})\,\mathbf{u}^*(\mathbf{m}) = \mathbf{f}^*$$
$$\mathbf{d}^* = \mathbf{s}^*\mathbf{u}^* \tag{129}$$

We take the derivative of equation 125 and 129 with respect to m_l, we get

$$\frac{\partial}{\partial m_l}(\mathbf{H}\mathbf{u}) = \frac{\partial \mathbf{f}}{\partial m_l}$$
$$\frac{\partial}{\partial m_l}(\mathbf{H}^*\mathbf{u}^*) = \frac{\partial \mathbf{f}^*}{\partial m_l} \tag{130}$$

As the source signature is independent of model parameter $\mathbf{m}$, the expression shows in equation 130 can be simplified as

$$\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u} + \mathbf{H}\frac{\partial \mathbf{u}}{\partial m_l} = \mathbf{0}$$
$$\frac{\partial \mathbf{H}^*}{\partial m_l}\mathbf{u}^* + \mathbf{H}^*\frac{\partial \mathbf{u}^*}{\partial m_l} = \mathbf{0} \tag{131}$$

So we can get the expression of $\frac{\partial \mathbf{u}}{\partial m_l}$ and $\frac{\partial \mathbf{u}^*}{\partial m_l}$, which are given as

$$\frac{\partial \mathbf{u}}{\partial m_l} = -\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}$$
$$\frac{\partial \mathbf{u}^*}{\partial m_l} = -(\mathbf{H}^*)^{-1}\frac{\partial \mathbf{H}^*}{\partial m_l}\mathbf{u}^* \tag{132}$$

The next step we need to figure out the expression of $\frac{\partial J}{\partial \mathbf{u}}$ and $\frac{\partial J}{\partial \mathbf{u}^*}$. As equation 127 can be also written as

$$J = \frac{1}{2}(\mathbf{s}\mathbf{u} - \mathbf{d}_{obs})^T(\mathbf{s}\mathbf{u}^* - \mathbf{d}_{obs}^*) \tag{133}$$

It is easy to derive that

$$\frac{\partial J}{\partial \mathbf{u}} = \frac{1}{2}\mathbf{s}^T(\mathbf{d}^* - \mathbf{d}_{obs}^*)$$
$$\frac{\partial J}{\partial \mathbf{u}^*} = \frac{1}{2}\mathbf{s}^T(\mathbf{d} - \mathbf{d}_{obs}) \tag{134}$$

We combine the result of equation 128, 132 and 134, we can get

$$
\begin{aligned}
\frac{\partial J}{\partial m_l} &= -\frac{1}{2}\left(\left(\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T \mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs}) + \left((\mathbf{H}^*)^{-1}\frac{\partial \mathbf{H}^*}{\partial m_l}\mathbf{u}^*\right)^T \mathbf{s}^T(\mathbf{d} - \mathbf{d}_{obs})\right) \\
&= -\frac{1}{2}\left(\left(\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T \mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs}) + \left(\left(\mathbf{H}^{-1}\right)^*\frac{\partial \mathbf{H}^*}{\partial m_l}\mathbf{u}^*\right)^T \mathbf{s}^T(\mathbf{d} - \mathbf{d}_{obs})\right) \quad (135) \\
&= -\frac{1}{2}\left(\left(\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T \mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs}) + \left(\left(\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T\right)^* \mathbf{s}^T(\mathbf{d} - \mathbf{d}_{obs})\right)
\end{aligned}
$$

It is not easily to find that the second term is the complex conjugate of the first term in the above equation. So the final form of the gradient can be simplified as

$$
\begin{aligned}
\frac{\partial J}{\partial m_l} &= -\mathcal{R}\left\{\left(\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T \mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs})\right\} \\
&= -\mathcal{R}\left\{\left(\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)^T (\mathbf{H}^T)^{-1}\mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs})\right\}
\end{aligned}
\quad (136)
$$

where $\mathcal{R}$ is the operator which take the real part of complex number. People usually call $\left(\frac{\partial \mathbf{H}}{\partial m_l}\mathbf{u}\right)$ as the source side wave field and $(\mathbf{H}^T)^{-1}\mathbf{s}^T(\mathbf{d}^* - \mathbf{d}^*_{obs})$ as adjoint wave field, the gradient is obtained by the inner product between source-side wave field and adjoint wave field.