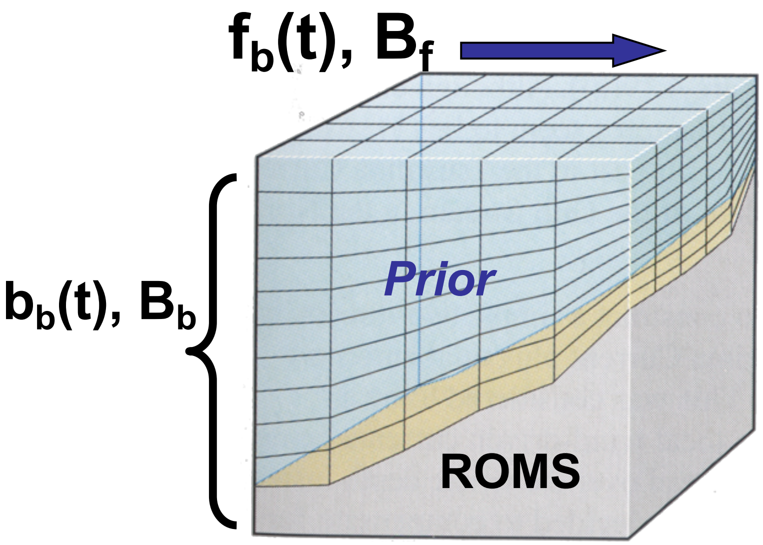
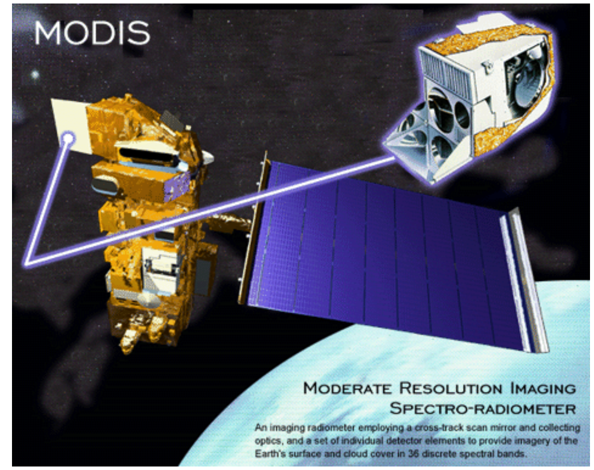
Lecture 3:
Dual 4D-Var

Outline

- 4D-Var recap
- Dual 4D-Var
- The ROMS dual algorithms
- Weak constraint 4D-Var

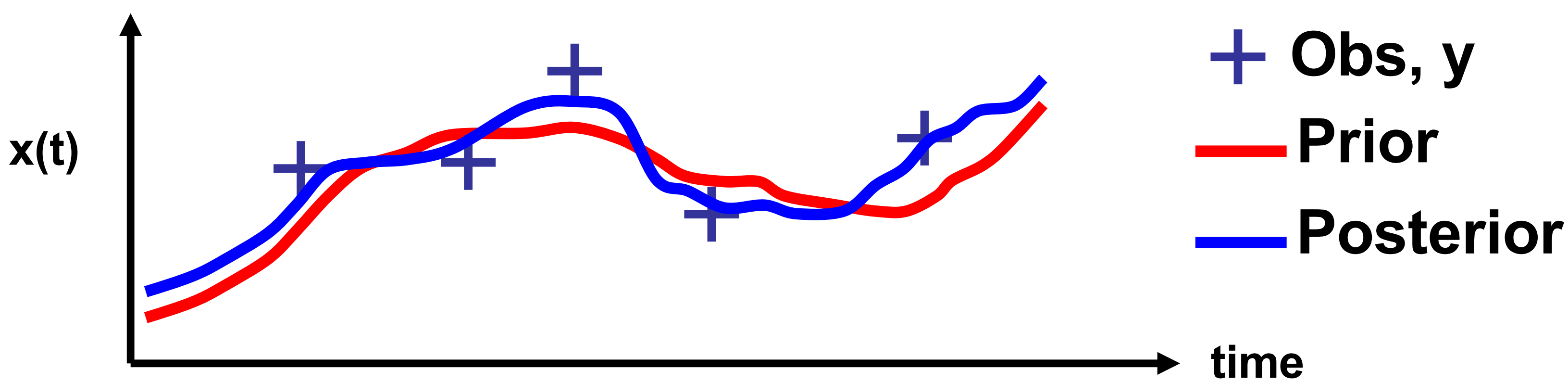
Data Assimilation: Recap





$x_b(0)$, B_x

y, R



Model solutions depends on $x_b(0)$, $f_b(t)$, $b_b(t)$, $h(t)$

Notation & Nomenclature: Recap

$$\mathbf{x} = \begin{bmatrix} \mathbf{T} \\ \mathbf{S} \\ \boldsymbol{\zeta} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix}$$

State vector

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{f}(t) \\ \mathbf{b}(t) \\ \boldsymbol{\eta}(t) \end{bmatrix}$$

Control vector

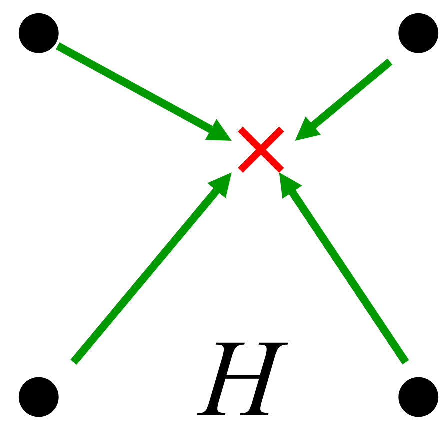
$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_N \end{bmatrix}$$

Observation vector

Prior

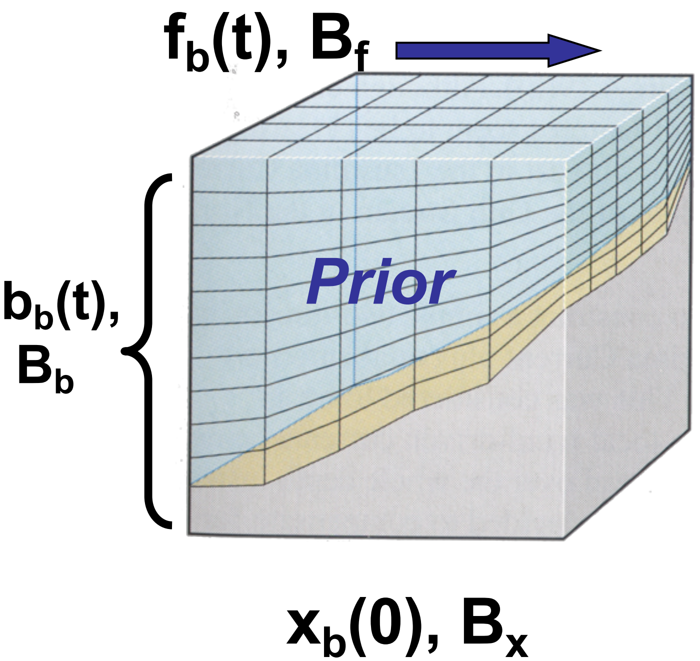
$$\mathbf{d} = \left(\mathbf{y} - H(\mathbf{x}^{\mathrm{b}})\right)$$

Innovation vector



Observation operator

Incremental Formulation: Recap



$$\delta \mathbf{z} = (\delta \mathbf{x}^T(0), \boldsymbol{\varepsilon}_{\mathbf{b}}^T(t), \boldsymbol{\varepsilon}_{\mathbf{f}}^T(t), \boldsymbol{\eta}(t))^T$$

initial condition increment

boundary condition increment

forcing increment

corrections for model error

(Courtier *et al.*, 1994)

$$J = \frac{1}{2}\delta \mathbf{z}^T \mathbf{D}^{-1} \delta \mathbf{z} + \frac{1}{2}(\mathbf{G}\delta \mathbf{z} - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{G}\delta \mathbf{z} - \mathbf{d})$$

$$\mathbf{D} = \underbrace{\mathrm{diag}(\mathbf{B}_{\mathbf{x}}, \mathbf{B}_{\mathbf{b}}, \mathbf{B}_{\mathbf{f}}, \mathbf{Q})}$$

Prior (background) error covariance

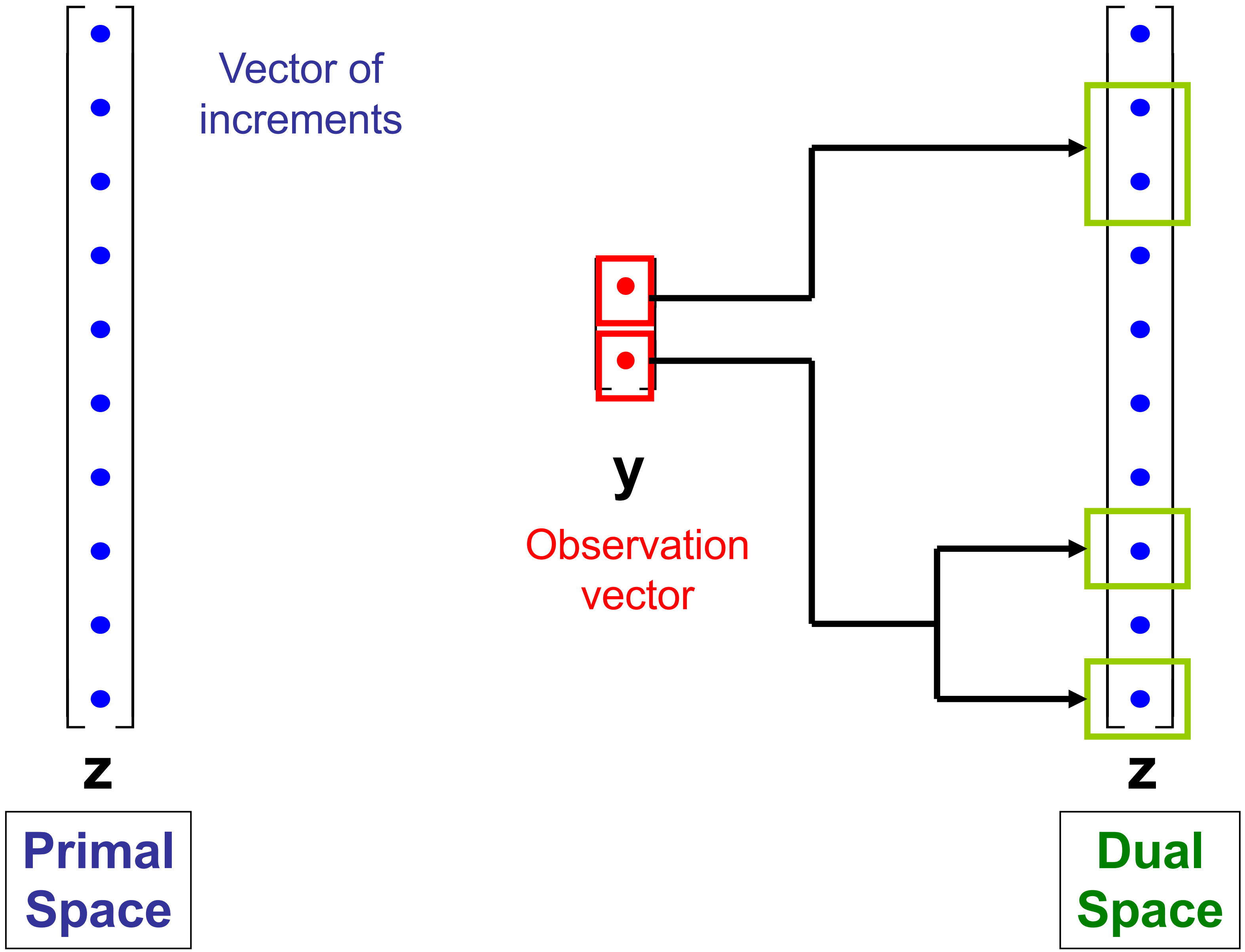
Tangent Linear Model sampled at obs points

Obs Error Cov.

Innovation

$$\mathbf{d} = \mathbf{y} - \mathbf{H}\mathbf{x}_{\mathrm{b}}$$

Primal vs Dual Formulation: Recap



The Solution: Recap

Analysis: $$\mathbf{z_a} = \mathbf{z_b} + \mathbf{Kd}$$

Gain (dual form):

$$\mathbf{K} = \mathbf{D}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}$$

Gain (primal form):

$$\mathbf{K} = (\mathbf{D}^{-1} + \mathbf{G}^{T}\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^{T}\mathbf{R}^{-1}$$

Two Spaces: Recap

Gain (dual):

$$\mathbf{K} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\underbrace{(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})}_{N_{\mathrm{obs}} \times N_{\mathrm{obs}}}{}^{-1}$$

Gain (primal):

$$\mathbf{K} = \underbrace{(\mathbf{D}^{-1} + \mathbf{G}^{T}\mathbf{R}^{-1}\mathbf{G})}_{N_{\mathrm{model}} \times N_{\mathrm{model}}}{}^{-1}\mathbf{G}^{T}\mathbf{R}^{-1}$$

$$\boxed{N_{\mathrm{obs}} \ll N_{\mathrm{model}}}$$

Two Spaces: Recap

G maps from model space
to observation space

$\mathbf{G}^T$ maps from observation space
to model space

Primal Formulation: Recap

Analysis: $\mathbf{z_a} = \mathbf{z_b} + \mathbf{Kd}$

Goal of 4D-Var is to identify:

$$\delta\mathbf{z} = \mathbf{Kd} = (\mathbf{D}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{d}$$

Solve the equivalent linear system:

$$(\mathbf{D}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G})\delta\mathbf{z} = \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{d}$$

by minimizing:

$$J = \frac{1}{2}\delta\mathbf{z}^T(\mathbf{D}^{-1} + \mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{G})\delta\mathbf{z} - \delta\mathbf{z}^T\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{R}^{-1}\mathbf{d}$$

$$= \frac{1}{2}\delta\mathbf{z}^T\mathbf{D}^{-1}\delta\mathbf{z} + \frac{1}{2}\left(\mathbf{G}\delta\mathbf{z} - \mathbf{d}\right)^T\mathbf{R}^{-1}\left(\mathbf{G}\delta\mathbf{z} - \mathbf{d}\right)$$

Dual Formulation

Analysis: $$\mathbf{z}_{\mathbf{a}} = \mathbf{z}_{\mathbf{b}} + \mathbf{K}\mathbf{d}$$

Goal of 4D-Var is to identify:

$$\delta\mathbf{z} = \mathbf{K}\mathbf{d} = \mathbf{D}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}$$

Solve the equivalent linear system:

$$(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})\mathbf{w} = \mathbf{d}; \qquad \delta\mathbf{z} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{w}$$

by minimizing:

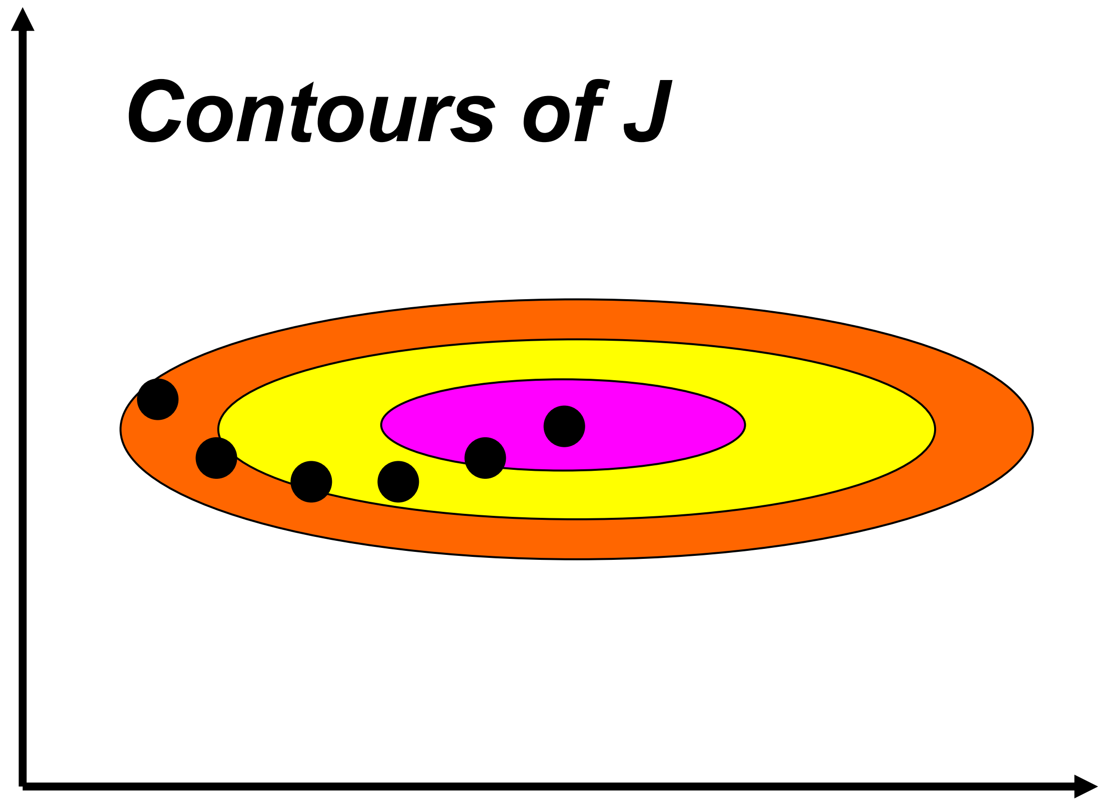
There is no physical significance attached to w

$$I(\mathbf{w}) = \frac{1}{2}\mathbf{w}^{T}(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})\mathbf{w} - \mathbf{w}^{T}\mathbf{d}$$

then compute:

$$\delta\mathbf{z} = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{w}$$

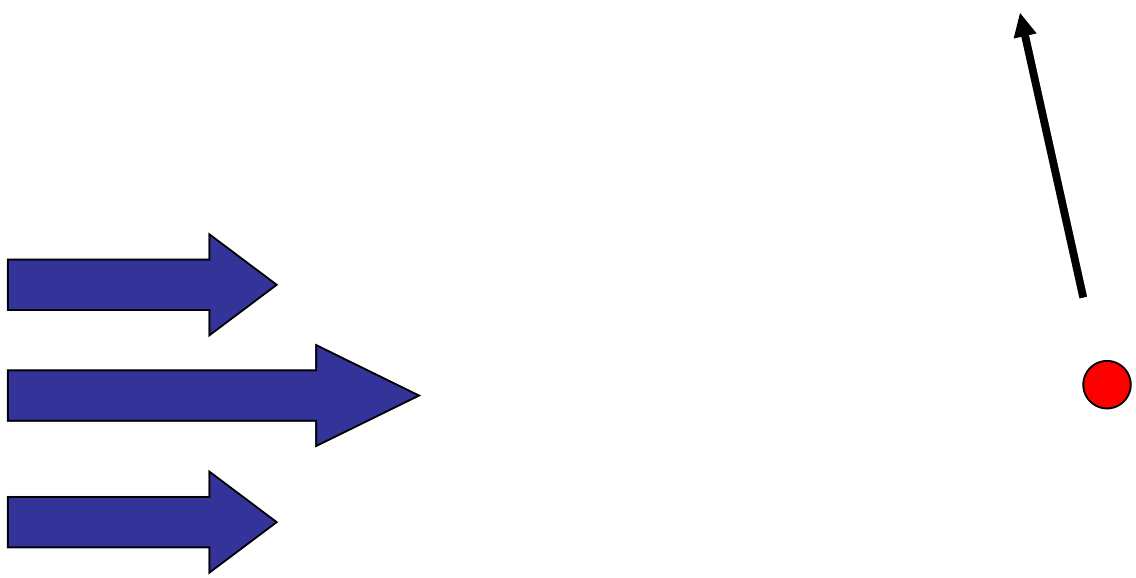
Conjugate Gradient (CG) Methods



Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{GDG}^{\mathrm{T}}\boldsymbol{\delta}$$



Zonal shear flow

Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{G}\mathbf{D}\mathbf{G}^{T}\delta$$

Adjoint Model

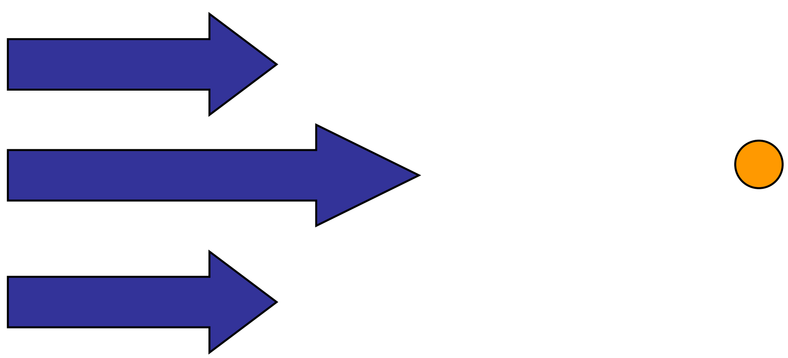
Zonal shear flow

Matrix-less Operations

There are no matrix multiplications!

$$GDG^{T}\delta$$

Adjoint Model



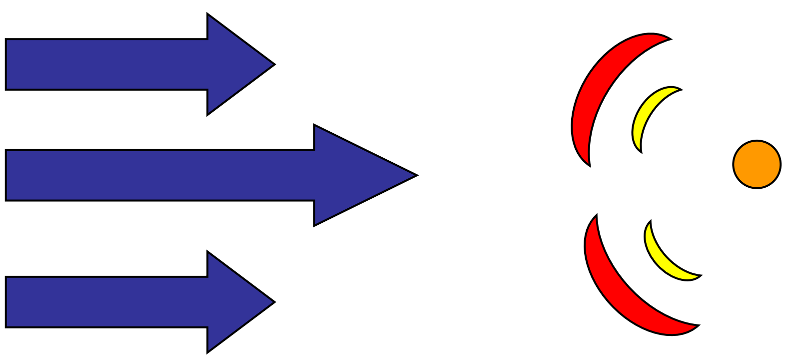
Zonal shear flow

Matrix-less Operations

There are no matrix multiplications!

$$GDG^{T}\delta$$

Adjoint Model



Zonal shear flow

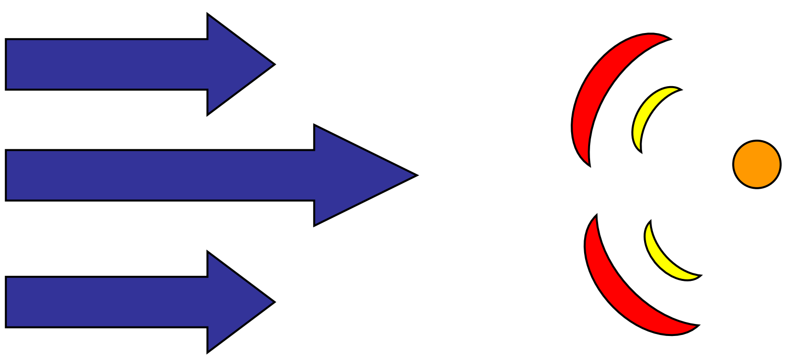
Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{G D G}^{T} \boldsymbol{\delta}$$

$$\uparrow$$

Covariance



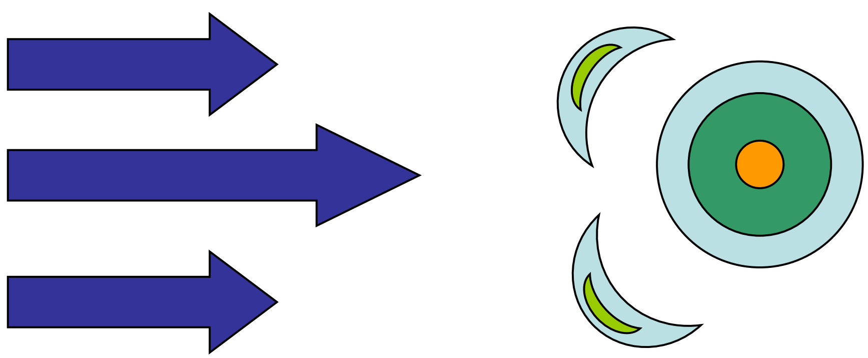
Zonal shear flow

Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{G}\mathbf{D}\mathbf{G}^{T}\delta$$

Covariance



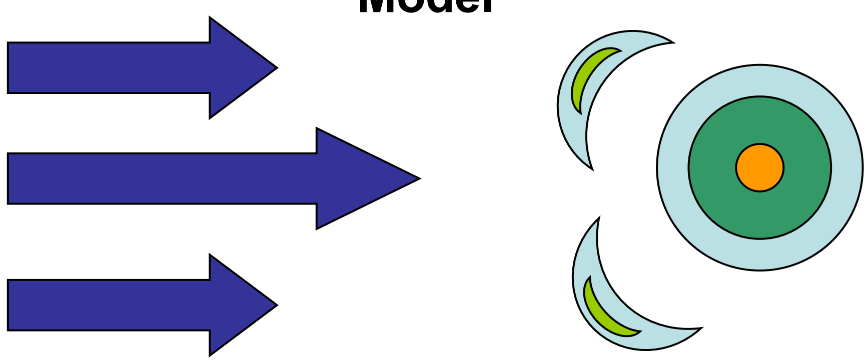
Zonal shear flow

Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{G D G}^{T} \boldsymbol{\delta}$$

Tangent Linear Model



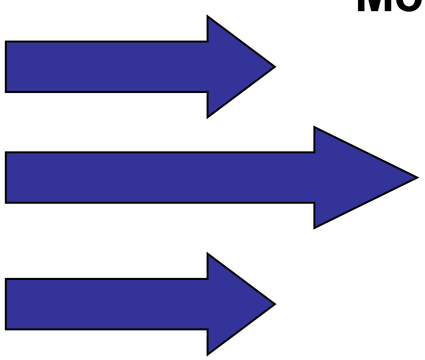
Zonal shear flow

Matrix-less Operations

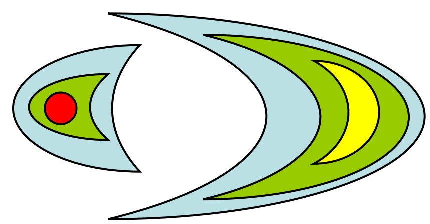
There are no matrix multiplications!

$$GDG^T\delta$$

Tangent Linear Model



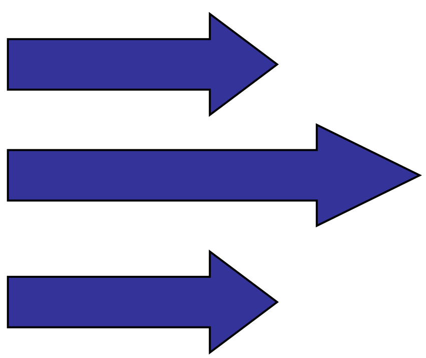
Zonal shear flow



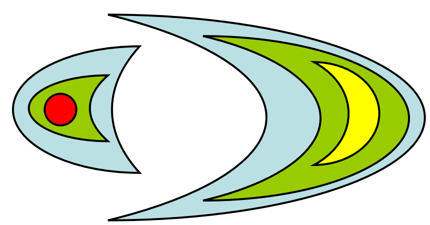
Matrix-less Operations

There are no matrix multiplications!

$$GDG^T\delta$$



Zonal shear flow



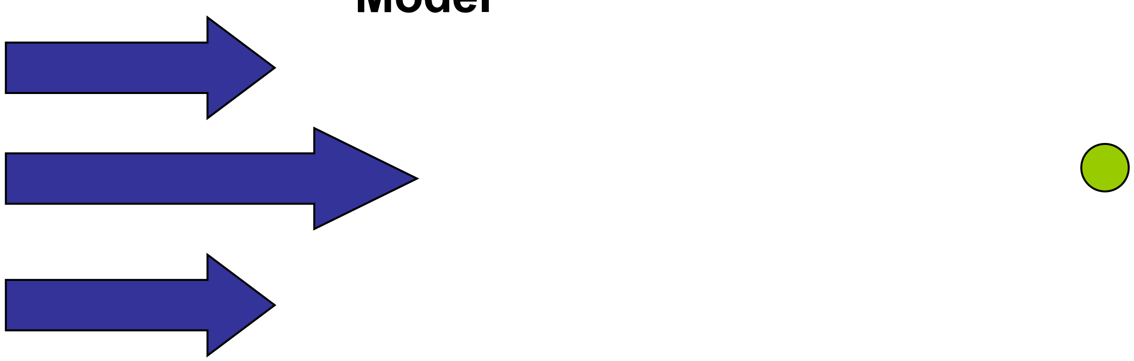
A covariance

Matrix-less Operations

There are no matrix multiplications!

$$\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}}\boldsymbol{\delta}$$

Tangent Linear Model

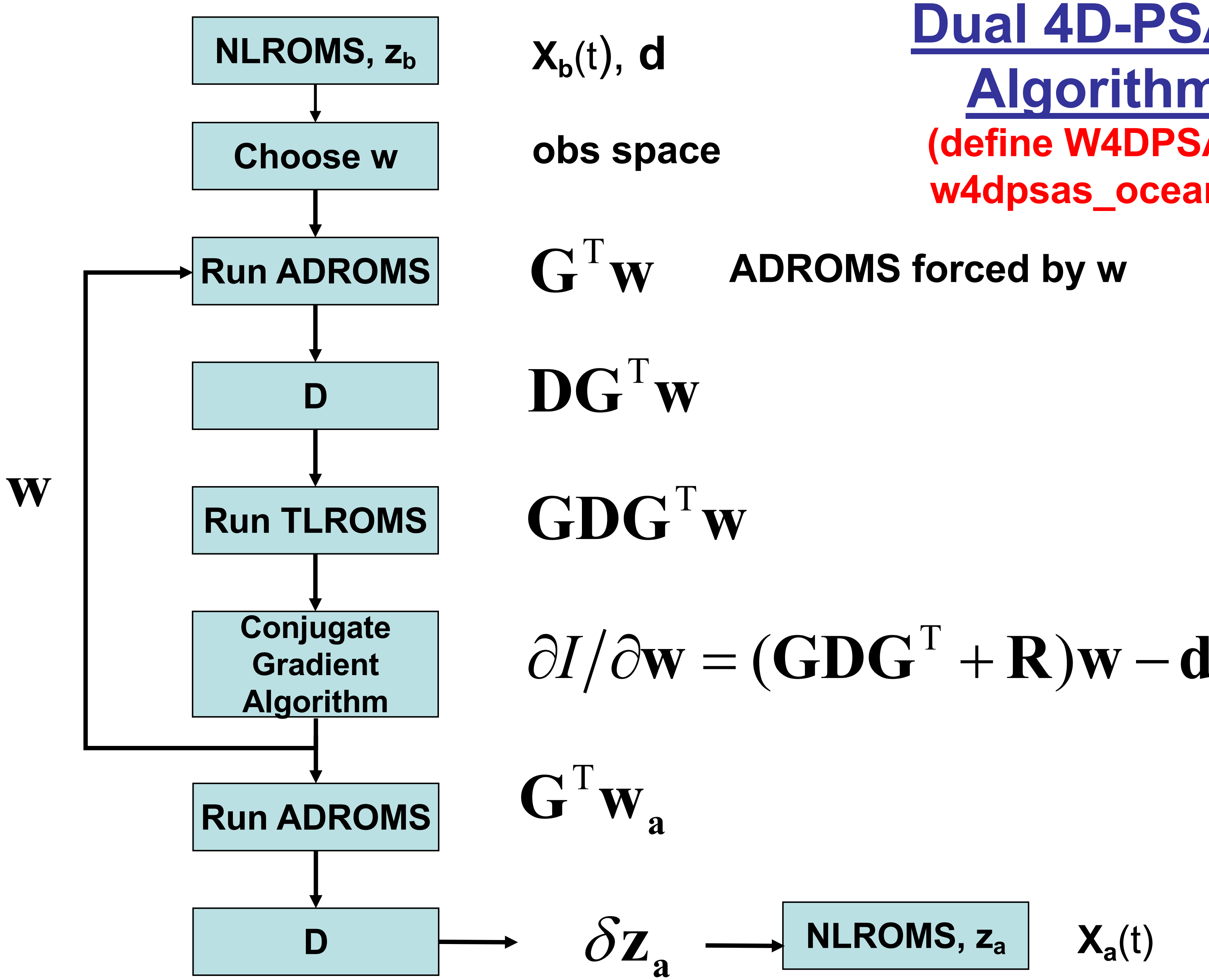


Zonal shear flow

Physical-space Statistical Analysis System (PSAS) – Da Silva *et al.* (1995)

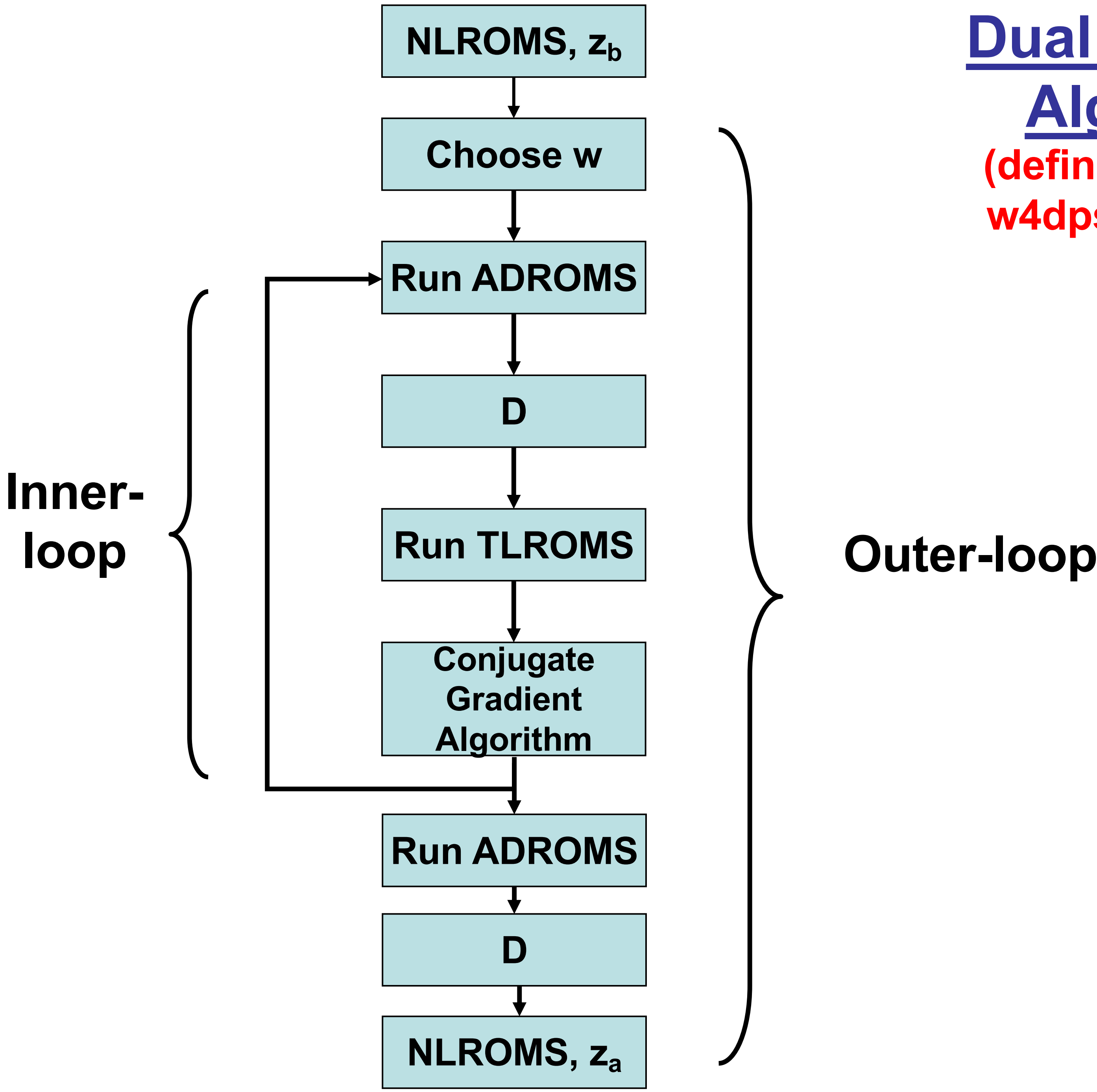
Dual 4D-PSAS Algorithm

(define W4DPSAS, w4dpsas_ocean.h)



Dual 4D-PSAS Algorithm

(define W4DPSAS, w4dpsas_ocean.h)



Weak Constraint 4D-Var

Nonlinear ROMS (NLROMS):

$$\mathbf{x}(t_i) = M(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i))$$

Nonlinear ROMS (NLROMS) with model error:

$$\mathbf{x}(t_i) = M(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i), \boldsymbol{\varepsilon}(t_i))$$

Model error *prior*: 0

Model error *prior* covariance: $\mathbf{Q}$

(no explicit time correlation in Q, but there is some in practice)

4D-Var control vector:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{f}(t) \\ \mathbf{b}(t) \\ \boxed{\boldsymbol{\eta}(t)} \end{bmatrix}$$

Correction for model error

Weak Constraint 4D-Var

Tangent linear ROMS (TLROMS):

$$\delta\mathbf{x}(t_i) = \mathbf{M}(t_i, t_{i-1})\delta\mathbf{u}(t_{i-1})$$

$$\delta\mathbf{u}(t_i) = \begin{bmatrix} \delta\mathbf{x}(t_i) \\ \delta\mathbf{f}(t_i) \\ \delta\mathbf{b}(t_i) \\ \boxed{\delta\boldsymbol{\eta}(t_i)} \end{bmatrix}$$

4D forcing for TLROMS

Strong constraint: $\delta\boldsymbol{\eta}(t_i) = \mathbf{0}$

Two Spaces

Gain (dual):

$$\mathbf{K} = \mathbf{D}\mathbf{G}^{\mathrm{T}}(\underbrace{\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R}}_{N_{\mathrm{obs}} \times N_{\mathrm{obs}}})^{-1}$$

Gain (primal):

$$\mathbf{K} = (\underbrace{\mathbf{D}^{-1} + \mathbf{G}^{T}\mathbf{R}^{-1}\mathbf{G}}_{N_{\mathrm{model}} \times N_{\mathrm{model}}})^{-1}\mathbf{G}^{T}\mathbf{R}^{-1}$$

$$\boxed{N_{\mathrm{obs}} \ll N_{\mathrm{model}}}$$

Two Spaces

Strong constraint:

$$N_{\mathrm{model}} = N_x + N_{times}\left(N_f + N_b\right)$$

Weak constraint:

$$N_{\mathrm{model}} = N_x + N_{times}\left(N_f + N_b + \boxed{N_x}\right)$$

Weak constraint is only practical in dual formulation of 4D-Var since *N_{obs}* is unaffected:

$$\mathbf{K} = \mathbf{DG}^{\mathrm{T}}\underbrace{(\mathbf{GDG}^{\mathrm{T}} + \mathbf{R})}_{N_{\mathrm{obs}} \times N_{\mathrm{obs}}}^{-1}$$

Dual 4D-Var & Preconditioning

- While it appears that the dual formulation should be an easier problem to solve because of the considerably smaller dimension of the space involved, it has not been widely adopted because of practical barriers to convergence.
- Until recently preconditioning was a major hurdle to advances in the dual algorithm.
- We will review some recent progress in the framework of ROMS:
  (i) $\mathbf{R}^{-1/2}$ preconditioning (PSAS – Courtier, 1997)
  (ii) MINRES vs CG (El Akkraoui and Gauthier, 2010)
  (iii) RPCG (Gratton and Tshimanga, 2009)

Dual 4D-Var: Naïve $R^{-1/2}$ Preconditioning

Analysis: $$\mathbf{z_a} = \mathbf{z_b} + \mathbf{Kd}$$

Goal of 4D-Var is to identify:

$$\delta\mathbf{z} = \mathbf{Kd} = \mathbf{DG}^{\mathrm{T}}(\mathbf{GDG}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}$$

Solve the equivalent linear system:

$$(\mathbf{GDG}^{\mathrm{T}} + \mathbf{R})\mathbf{w} = \mathbf{d}; \qquad \delta\mathbf{z} = \mathbf{DG}^{\mathrm{T}}\mathbf{w}$$

by minimizing:

$$I(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T(\mathbf{GDG}^{\mathrm{T}} + \mathbf{R})\mathbf{w} - \mathbf{w}^T\mathbf{d}$$

Preconditioning via the change of variable

$$\mathbf{v} = \mathbf{R}^{-1/2}\mathbf{w}$$

Dual 4D-Var: CG with $R^{-1/2}$ Preconditioning

Lanczos formulation of conjugate gradient algorithm in observation space is used (congrad.F).

Dual formulation of gain matrix:

$$\mathbf{K} = \mathbf{D}\mathbf{G}^{\mathrm{T}}(\mathbf{G}\mathbf{D}\mathbf{G}^{\mathrm{T}} + \mathbf{R})^{-1}$$

Dual formulation of practical gain matrix:

$$\tilde{\mathbf{K}}_k = \mathbf{D}\mathbf{G}^{\mathrm{T}}\mathbf{R}^{-1/2}\tilde{\mathbf{V}}_k\mathbf{T}_k^{-1}\tilde{\mathbf{V}}_k^{\mathrm{T}}\mathbf{R}^{-1/2}$$

$\tilde{\mathbf{V}}_k$ **dual Lanczos vectors**

Dual 4D-Var: Minimum Residual Method

Lanczos formulation of minimum residual algorithm in observation space is used (define MINRES).

Dual 4D-Var: Restricted Preconditioned CG (RPCG)

- **Experience in primal 4D-Var has shown that preconditioning by $B^{-1/2}$ is very effective.**
- **Preconditioning by $B^{-1/2}$ can be enforced in dual 4D-Var by "restricting" the dual Lanczos vectors so that:**

$$\underbrace{\mathbf{V}_k}_{\text{Primal Lanczos vectors}} = \mathbf{G}^{\mathrm{T}} \underbrace{\tilde{\mathbf{V}}_k}_{\text{Dual Lanczos vectors}}$$

- **define RPCG**

An Example: ROMS CCS

COAMPS forcing

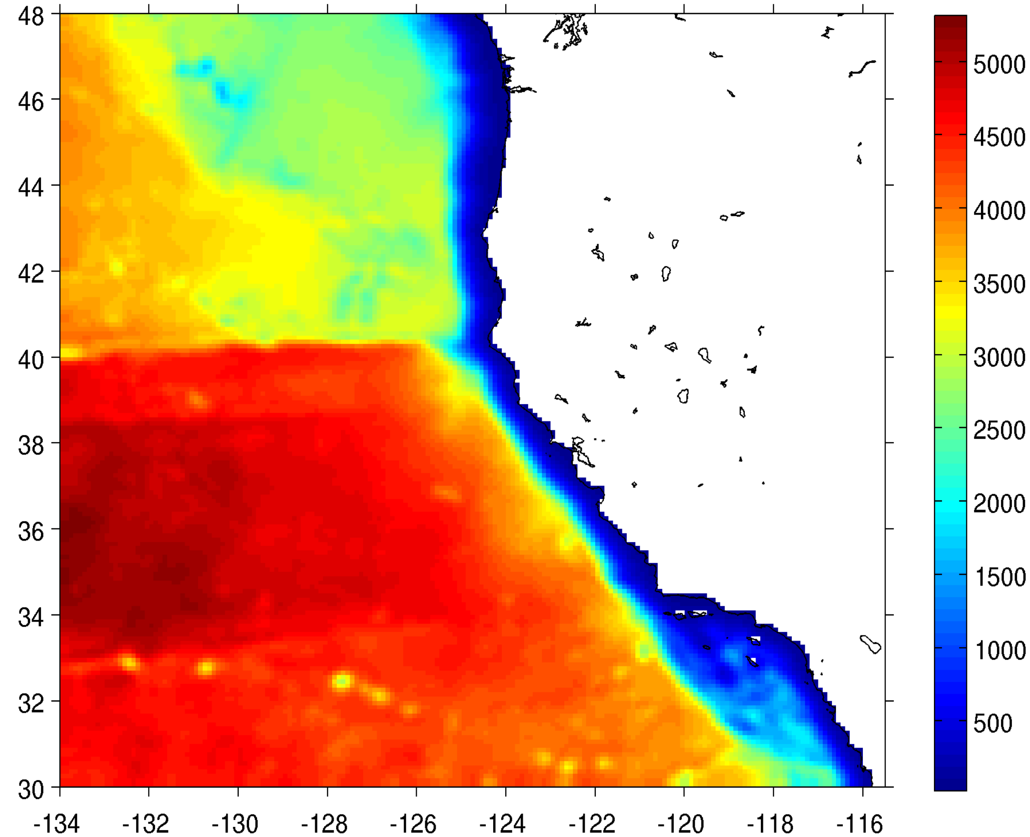
$f_b(t), B_f$

ECCO open boundary conditions

$b_b(t), B_b$

$x_b(0), B_x$

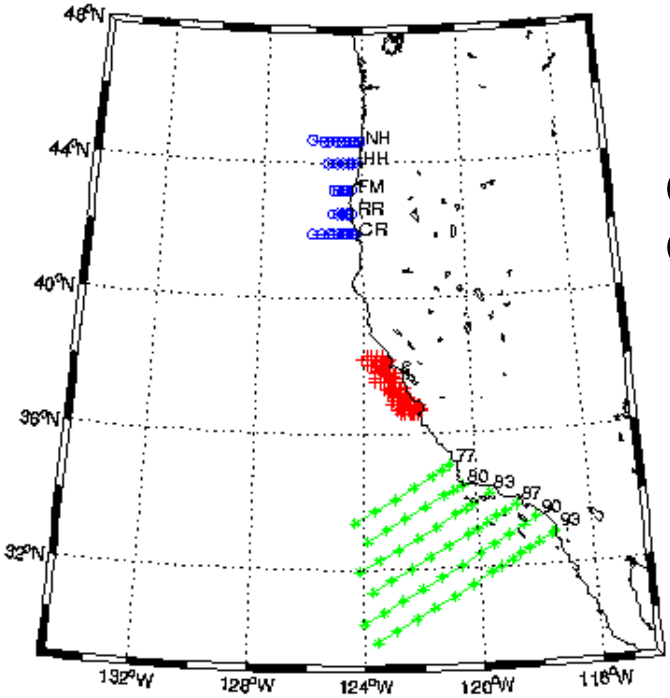
Previous assimilation cycle



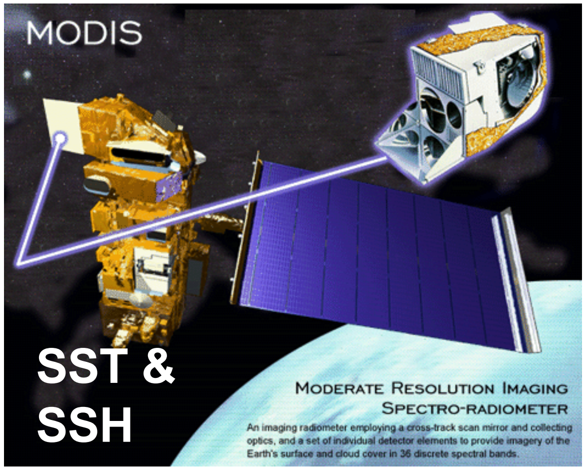
30km, 10 km & 3 km grids, 30- 42 levels

Veneziani et al (2009)
Broquet et al (2009)
Moore et al (2010)

Observations (y)



CalCOFI & GLOBEC



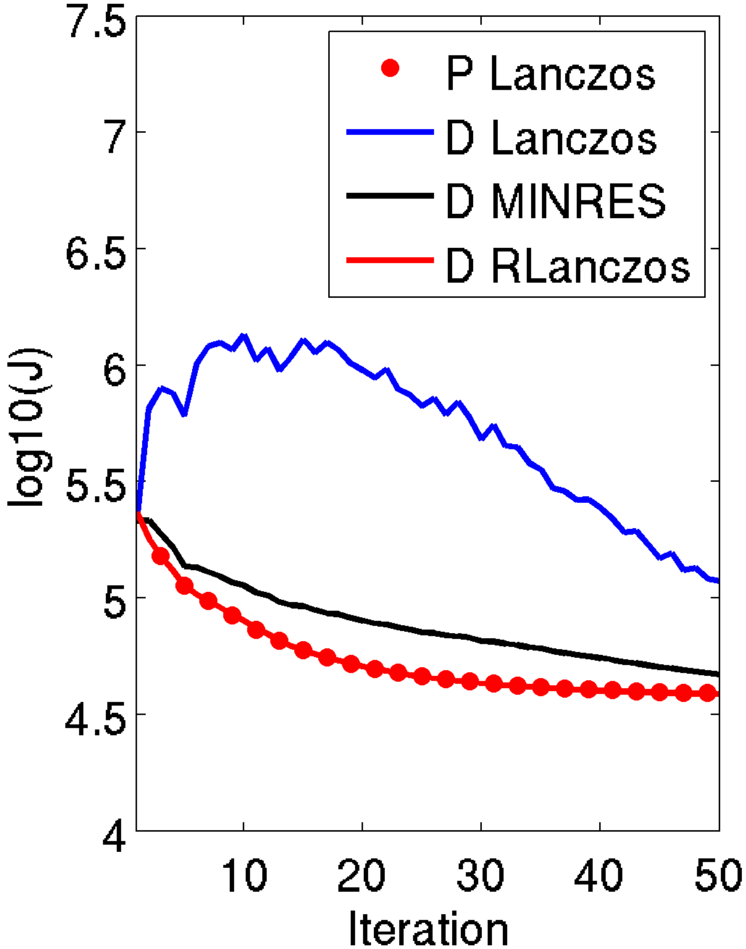
Ingleby and Huddleston (2007)

4D-Var Configuration

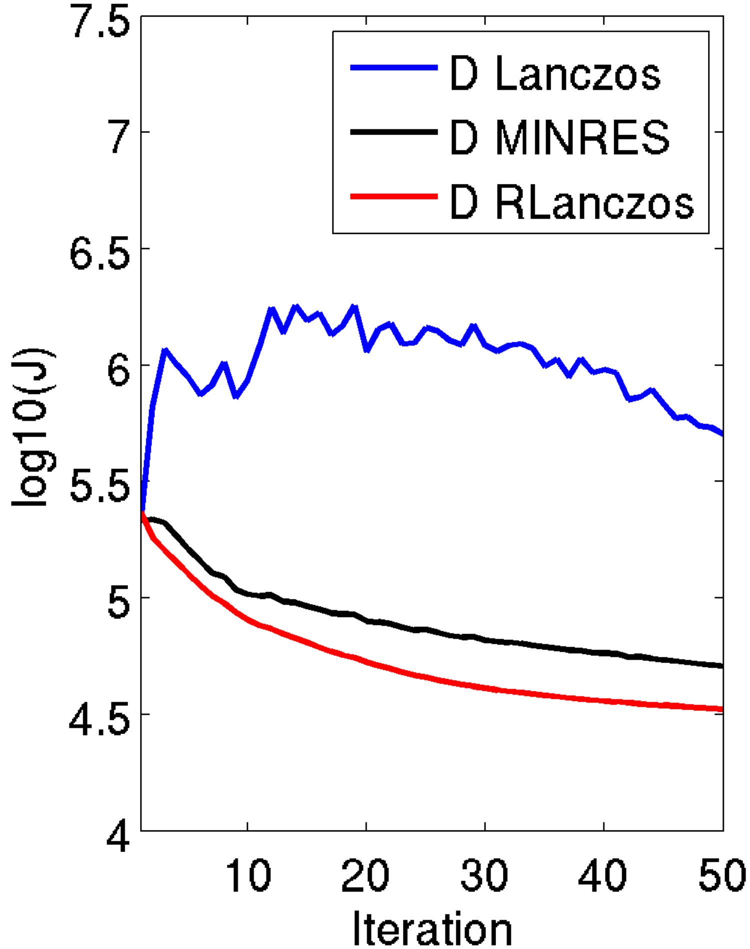
- Case studies for a representative case
  29 March - 4 April, 2003, 10km resolution.
- 1 outer-loop, 50 inner-loops
- 7 day assimilation window
- *Prior* **D***:* **x** L_h=50 km, L_v=30m, σ from clim
  f L_τ=300km, L_Q=100km, σ from COAMPS
  b L_h=100 km, L_v=30m, σ from clim
- Super observations formed
- Obs error **R** (diagonal):
  SSH 2 cm
  SST 0.4 C
  hydrographic 0.1 C, 0.01psu

Dual 4D-Var

(Gürol et al, 2013, QJRMS)

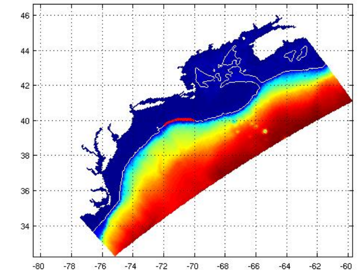


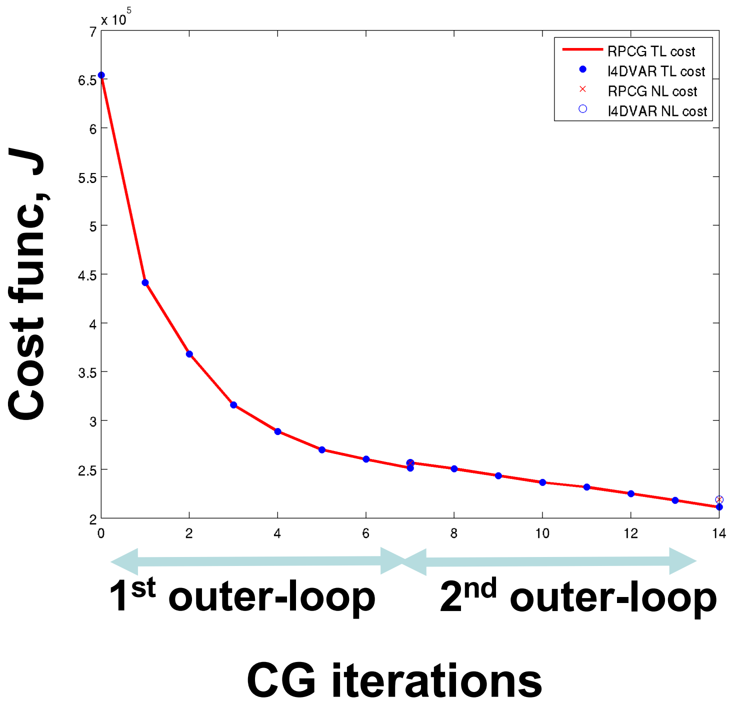
Strong Constraint



Weak Constraint

Dual 4D-Var vs Primal 4D-Var





Mid-Atlantic Bight, 7km, 2014 test case
2 outer-loops
7 inner-loops
3 day windows
SSH, SST, in situ, HF radar obs

- **Dual 4D-Var is ~15-25% faster than primal 4D-Var in ROMS**
- **Dual 4D-Var has more post-processing utility in ROMS**
- **Only Dual 4D-Var supports the weak constraint option**

Summary

- Strong and weak constraint 4D-Var, dual formulation:

  define W4DPSAS

  Drivers/w4dpsas_ocean.h

- Matrix-less iterations to identify cost function minimum using TLROMS and ADROMS

- RPCG: define RPCG

References

- Bennett, A.F., 2002: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, 234pp.
- Courtier, P., 1997: Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteorol. Soc.*,**123**, 2449-2461.
- Da Silva, A., J. Pfaendtner, J. Guo, M. Sienkiewicz and S. Cohn, 1995: Assessing the effects of data selection with DAO's physical-space statistical analysis system. Proceedings of the second international WMO symposium on assimilation of observations in meteorology and oceanography, Tokyo 13-17 March, 1995. WMO.TD 651, 273-278.
- Di Lorenzo, E., A.M. Moore, H.G. Arango, B.D. Cornuelle, A.J. Miller, B. Powell, B.S. Chua and A.F. Bennett, 2007: Weak and strong constraint data assimilation in the inverse Regional Ocean Modeling System (ROMS): development and application for baroclinic coastal upwelling system. *Ocean Modeling*, **16**, 160-187.
- Egbert, G.D., A.F. Bennett and M.C.G. Foreman, 1994: TOPEX/POSEIDON tides estimated using a global inverse method. *J. Geophys. Res.*, **99**, 24,821-24,852.
- El Akkraoui, A. and P. Gauthier, 2010: Convergence properties of the primal and dual forms of variational data assimilation. *Q. J. R. Meteorol. Soc.*, **136**, 107-115.

References

- Gratton, S., and J.Tshimanga, 2009: An observation-space formulation of variational assimilation using a restricted preconditioned conjugate gradient algorithm. *Q. J. R. Meteorol. Soc.*, **135**, 1573-1585.
- Gürol, S., S. Gratton, A.T. Weaver, A.M. Moore, A. Piacentini and H.G. Arango, 2014: Minimization algorithms for variational data assimilation with the dual formulation. *Q. J. Roy. Meteorol. Soc.*, **140**, 539-556.
- Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, J. Zavala-Garay and A.T. Weaver, 2011: The Regional Ocean Modeling System (ROMS) 4-dimensional data assimilation systems. I: System overview and formulation. *Prog. Oceanogr.*, **91**, 34-49.