



**Thèse de doctorat d'Université
Paris-Est**

Sciences, Ingénierie et Environnement

Présentée et soutenue par:

Anthony Fillion

Méthodes variationnelles d'ensemble et
optimisation variationnelle pour les
géosciences

Directeurs de thèse :

Marc Bocquet
Serge Gratton

CEREA, École des Ponts Paris Tech
IRIT, INP Toulouse

Rapporteurs :

Étienne Mémin
Yann Michel

FLUMINANCE, INRIA Rennes
GMAP, Météo-France Toulouse

Examineurs :

Maëlle Nodet
Selime Gürol
Annick Sartenaer

AIRSEA, INRIA Grenoble
CERFACS, Toulouse
NaXys, Université de Namur

Abstract

L'assimilation de données consiste à calculer une estimation de l'état d'un système physique. Cette estimation doit alors combiner de façon optimale des observations entachées d'erreurs de mesure et des modèles numériques imparfaits permettant de simuler le système physique. En pratique, l'assimilation de données sert à estimer l'état initial d'un système dynamique. Cet état analysé peut ensuite être utilisé pour prévoir le comportement de ce système, notamment dans les systèmes géophysiques où les jeux de données sont conséquents.

Une première approche repose sur une estimation de l'état initial basée sur le principe du maximum a posteriori. Il s'agit alors de résoudre un problème d'optimisation, souvent par des techniques utilisant le gradient des opérateurs. Cette approche, appelée 4DVar, nécessite le calcul de l'adjoint du modèle et de l'opérateur d'observation, ce qui est une tâche consommatrice en temps de développement des systèmes de prévision. Une seconde approche permettant de résoudre séquentiellement le problème d'assimilation est basée sur les techniques dites « d'ensemble ». Ici, des perturbations a priori de l'état du système permettent d'estimer des statistiques. Ces moments sont alors utilisés dans les formules de Kalman pour obtenir des approximations de l'état du système a posteriori.

Ces deux approches ont été récemment combinées avec succès dans les méthodes de type EnVar aujourd'hui utilisées dans les systèmes opérationnels de prévision. Elles bénéficient donc d'une gestion efficace de la non linéarité au travers des méthodes d'optimisation variationnelle et permettent l'estimation de statistiques et de dérivées à l'aide des ensembles. L'IEEnKS ([Bocquet and Sakov, 2014](#)) est un archétype de ces méthodes EnVar. Pour combiner les deux approches précédentes, il utilise une fenêtre d'assimilation qui est translatée entre chaque cycle. Différents paramétrages de la fenêtre d'assimilation conduisent à différentes stratégies d'assimilation non équivalentes lorsque la dynamique du système est non linéaire.

En particulier, les longues fenêtres d'assimilation réduisent la fréquence de l'approximation Gaussienne des densités a priori. Il en résulte une amélioration des performances jusqu'à un certain point. Au delà, la complexité structurelle de la fonction de coût met l'analyse variationnelle en défaut. Une solution nommée "quasi statique variational as-

similation” (QSVA) (Pires et al., 1996) permet d’atténuer ces problèmes en ajoutant graduellement les observations à la fonction de coût du 4DVar. Le second chapitre de thèse provenant de Fillion et al. (2018) généralise cette technique aux méthodes EnVar et s’intéresse plus précisément aux aspects théoriques et numériques du QSVA appliqués à l’IEnKS.

Cependant, l’intérêt du QSVA repose sur la perfection du modèle pour simuler l’évolution de l’état. En effet, la pertinence d’une observation temporellement éloignée pour estimer l’état peut être remise en cause en présence d’erreur modèle. Le troisième chapitre est consacré à l’introduction d’erreur modèle au sein de l’IEnKS en poursuivant les travaux de Sakov and Bocquet (2018); Sakov et al. (2018). Il y sera donc construit l’IEnKS-Q, une méthode 4D variationnelle d’ensemble résolvant séquentiellement le problème de lissage en présence d’erreur modèle. Malheureusement, en présence d’erreur modèle, une trajectoire n’est plus déterminée par son état initial. Le nombre de paramètres nécessaires à la caractérisation de ses statistiques augmente alors avec la longueur de la fenêtre d’assimilation. Lorsque ce nombre va de pair avec le nombre d’évaluations du modèle, les conséquences pour le temps de calcul sont catastrophiques. La solution proposée est alors de découpler ces quantités avec une décomposition des matrices d’anomalies. Dans ce cas, l’IEnKS-Q n’est pas plus coûteux que l’IEnKS en nombre d’évaluations du modèle.

Table des matières

Table des matières	v
Liste des symboles	vii
1 Introduction	1
1.1 L'assimilation de données sans dimension temporelle	1
1.1.1 Modélisation des probabilités	2
1.1.2 Solution bayésienne	5
1.1.3 Méthode par échantillonnage d'importance	6
1.1.4 Spécifications gaussiennes a priori	8
1.1.5 Échantillonner une densité	10
1.1.6 Solution dans le cadre linéaire et gaussien	14
1.1.7 Méthode dans un cadre moins linéaire et gaussien	18
1.1.8 Méthodes variationnelles	21
1.1.9 Conclusion de l'assimilation de données sans dimension temporelle	25
1.2 L'assimilation de données avec dimension temporelle	27
1.2.1 Solution bayésienne	27
1.2.2 Le filtre particulaire	31
1.2.3 Solution dans le cadre linéaire : le filtre de Kalman	33
1.2.4 Solution dans le cadre linéaire : le lisseur de Kalman	35
1.2.4.1 Remarque sur la méthode aller-retour	37
1.2.5 Le filtre de Kalman d'ensemble stochastique	38
1.2.5.1 Localisation et inflation	40
1.2.6 Le filtre de Kalman d'ensemble déterministe	41
1.2.7 Méthode variationnelle : 4DVar	44
1.2.8 Conclusion de l'assimilation de données avec dimension temporelle	45
1.3 L'assimilation de données ensemble variationnelle fortement contrainte .	47
1.3.1 Méthodes hybrides	47

1.3.2	Ensemble d'assimilations	48
1.3.3	Méthodes 4D variationnelles d'ensemble	48
1.3.4	Le lisseur de Kalman itératif	50
1.3.5	Conclusion de l'assimilation de données ensemble variationnelle fortement contrainte	56
2	Assimilation de données ensemble variationnelle quasi-statique	57
2.1	Introduction	57
2.1.1	Contexte	57
2.1.2	Assimilation de données variationnelle quasi-statique	58
2.1.3	Méthodes variationnelles d'ensemble	58
2.1.4	Plan	59
2.2	La fenêtre et les performances d'assimilation	60
2.2.1	Le 4DVar et l'IEEnKS	60
2.2.2	Performance de l'assimilation	64
2.2.3	Performance dans le cas linéaire, diagonal et autonome	66
2.2.4	Performance dans le cas non linéaire et chaotique	69
2.2.4.1	Des minimums locaux multiples	71
2.2.4.2	Longueur effective de la fenêtre d'assimilation	73
2.3	Algorithmes quasi-statiques	76
2.4	Expériences numériques	81
2.5	Conclusion	87
3	Le lisseur de Kalman itératif faiblement contraint	89
3.1	Introduction	89
3.1.1	Méthodes 4D ensemble variationnelles faiblement contraintes	89
3.1.2	Plan	90
3.2	Assimilation de données séquentielle faiblement contrainte dans l'espace de l'état	91
3.2.1	Le système d'assimilation de données	91
3.2.2	Principes d'assimilation de données séquentielle	92
3.2.3	Comparaison avec le biais d'erreur modèle comme variable de contrôle	96
3.3	Construction de l'IEEnKS-Q	98
3.3.1	Représentation dans le sous-espace de l'ensemble	98
3.3.2	Analyse	100
3.3.2.1	Analyse linéaire de la variable d'ensemble	100

3.3.2.2	Construction récursive des dérivées	101
3.3.2.3	Analyse variationnelle de la variable d'ensemble	102
3.3.2.4	Décomposition de la matrice de déviation et modèle de substitution	103
3.3.3	Propagation	105
3.3.3.1	Réduction de la dimension et marginalisation	106
3.4	Algorithmes	107
3.5	Résultats numériques	112
3.5.1	Validation de l'IEEnKS-Q	114
3.5.2	Test du paramètre G	116
3.5.3	Test avec des modèles de substitution	117
3.6	Conclusion	118
4	Conclusion	121
	References	127
A	Annexe de l'introduction	135
A.1	Maximum d'entropie	135
A.2	Minimum d'entropie relative	136
A.3	Meilleur estimateur en erreur quadratique moyenne	136
A.3.1	Avec E le sous-espace des fonctions affines	137
A.3.2	Avec E l'espace des fonctions	138
A.4	Variable d'ensemble et erreur d'échantillonnage	138
B	Annexe de l'assimilation de données ensemble variationnelle quasi-statique	141
B.1	Performance du 4D-Var dans le cas linéaire et uni-varié	141
B.2	Performance de l'IEEnKS dans le cas linéaire et uni-varié	143
B.3	Expression de la fonction de coût moyennée	145
C	Annexe du lisseur de Kalman itératif faiblement contraint	147
C.1	Incompatibilité entre les approches « biais d'erreur modèle » et « forçage d'erreur modèle »	147
C.2	Réduction de la fonction de coût 4DVar faiblement contrainte avec le biais d'erreur modèle comme variable de contrôle	148

Chapitre 1

Introduction

L'assimilation de données est une discipline visant à estimer l'état d'un système physique à l'aide d'observations physiques partielles. En géophysique, ce problème devient l'estimation des paramètres du système terrestre à l'aide de mesures terrestres bruitées et parcellaires. Lorsque les observations deviennent disponibles à mesure que le temps passe, il s'agit d'estimer séquentiellement la trajectoire du système à l'aide d'un historique d'observations. Le contexte géophysique impose un grand nombre de paramètres et d'observations atteignant les limites des ressources numériques disponibles ([Evensen, 2009](#)).

Dans ce chapitre d'introduction, nous présenterons le problème d'assimilation de données et discuterons de ses aspects théoriques et pratiques.

1.1 L'assimilation de données sans dimension temporelle

Le système géophysique est une abstraction des processus physiques terrestres. Il n'est donc pas la réalité mais une simplification que l'on espère suffisamment précise pour être capable de fournir de bonnes prévisions. Pour cela, l'état du système géophysique est quantifié par un vecteur réel représentant température, pression, humidité, etc en différents points du globe pour un total de plusieurs millions (voir milliards) de paramètres. On supposera donc que notre système est paramétrisé par un vecteur $\mathbf{x} \in \mathbb{R}^n$ avec n suffisamment grand pour ne pas pouvoir négliger les ressources numériques disponibles. De même, les observations physiques sont paramétrisées par un vecteur $\mathbf{y} \in \mathbb{R}^d$. Les équations régissant le système géophysique sont rassemblées dans l'opérateur d'observation \mathcal{H} . L'observation \mathbf{y} est donc sensée être déterminée par l'état \mathbf{x} avec une faible

erreur d'observation \mathbf{e} :

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \mathbf{e}. \quad (1.1)$$

C'est parce que le système n'est qu'une représentation simplifiée de la réalité dont les observations \mathbf{y} sont une mesure physique, que cette relation entre état et observation est inexacte.

Un objectif de l'assimilation de données consiste à déterminer l'état \mathbf{x} à partir des observations \mathbf{y} avec une faible erreur d'estimation \mathbf{n} :

$$\mathbf{x} = \mathcal{K}(\mathbf{y}) + \mathbf{n}. \quad (1.2)$$

Il s'agit d'un problème inverse ([Tarantola, 2004](#)). Malheureusement, sans plus d'informations sur l'erreur d'observation \mathbf{e} on ne peut déduire de \mathbf{y} aucune information sur \mathbf{x} .

On sait cependant qualitativement que l'opérateur d'observation est efficace : les petites erreurs sont plus probables que les grandes. Il faut alors quantifier l'aléatoire du système pour pouvoir tenir compte de cette information de nature probabiliste.

1.1.1 Modélisation des probabilités

L'objectif est ici d'associer à chaque réalisation du système une probabilité. Cette section introduira les variables aléatoires leurs densités de probabilité et leurs statistiques (consultez [Gallouët and Herbin \(2013\)](#) pour plus de détails sur la théorie de la mesure et de l'intégration et [Feller \(2008\)](#) pour une introduction à la théorie des probabilités).

Une variable est aléatoire lorsque sa valeur n'est pas déterminée. Les variables aléatoires sont à concevoir comme une fonction (mesurable) d'un espace inconnu Ω vers une puissance de \mathbb{R} . Par exemple \mathbf{x} sera la fonction :

$$\mathbf{x} : \begin{cases} \Omega & \rightarrow \mathbb{R}^n \\ \omega & \mapsto \mathbf{x} \end{cases}, \quad (1.3)$$

Le caractère aléatoire provient donc du fait qu'on ne sait pas quel ω est réalisé ou se réalisera. Ainsi, en travaillant avec \mathbf{x} , on travaille simultanément avec toutes ses réalisations \mathbf{x} possibles. Par exemple, dans les équations (1.1) et (1.2) les quantités $\mathbf{y}, \mathbf{x}, \mathbf{e}, \mathbf{n}$ se généralisent respectivement en les variables aléatoires $\mathbf{y}, \mathbf{x}, \boldsymbol{\varepsilon}, \boldsymbol{\eta}$:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}, \quad (1.4a)$$

$$\mathbf{x} = \mathcal{K}(\mathbf{y}) + \boldsymbol{\eta}. \quad (1.4b)$$

Une probabilité p est alors une mesure sur Ω et un évènement, une partie (mesurable) de Ω . On supposera que la probabilité d'un évènement du type " $\mathbf{x} \in E$ " avec $E \subseteq \mathbb{R}^n$ se mesure via l'intégrale :

$$p(\mathbf{x} \in E) = \int_E p(\mathbf{x} = \mathbf{x}) d\mathbf{x}, \quad (1.5)$$

où $p(\mathbf{x} = \mathbf{x})$ est la densité de probabilité de \mathbf{x} évaluée en \mathbf{x} ($d\mathbf{x}$ est la mesure de Lebesgue sur \mathbb{R}^n). Lorsqu'il n'y a pas d'ambiguïté, cette écriture sera simplifiée en :

$$p(\mathbf{x} \in E) = \int_E p(\mathbf{x}) d\mathbf{x}, \quad (1.6)$$

Les densités sont supposées normalisées :

$$p(\mathbf{x} \in \mathbb{R}^n) = \int_{\mathbb{R}^n} p(\mathbf{x}) d\mathbf{x} = 1, \quad (1.7a)$$

$$p(\mathbf{x} \in \emptyset) = \int_{\emptyset} p(\mathbf{x}) d\mathbf{x} = 0, \quad (1.7b)$$

de sorte que l'évènement certain a pour probabilité 1 et l'évènement impossible 0. Les densités de probabilités décrivent donc de manière analytique comment l'aléatoire d'une variable se distribue dans l'espace. L'aléatoire de \mathbf{x} est aussi caractérisée de manière statistique par l'espérance des fonctions de \mathbf{x} :

$$\mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})] = \int \varphi(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}, \quad (1.8)$$

pour toute fonction φ telle que l'expression précédente soit définie. Cette équation fait le lien entre les propriétés statistiques de la variable aléatoire \mathbf{x} et les propriétés analytiques de sa densité. En particulier, lorsque $p(\mathbf{x})$ est à décroissance rapide, on aura par intégration par partie (identité de Stein, [Johnson, 2004](#)) :

$$\mathbb{E}_{p(\mathbf{x})} [\nabla_{\mathbf{x}} \varphi(\mathbf{x})] = - \int \nabla_{\mathbf{x}} p(\mathbf{x}) \varphi(\mathbf{x})^T d\mathbf{x}, \quad (1.9)$$

avec $\nabla_{\mathbf{x}}$ le gradient suivant la variable \mathbf{x} et T la transposée. Lorsque φ est polynomiale,

les espérances de la forme $\mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})]$ sont les moments de \mathbf{x} . Nous nous intéresserons souvent aux moments d'ordre au plus deux. Ils comprennent la moyenne :

$$\mathbb{E}_{p(\mathbf{x})} [\mathbf{x}], \quad (1.10)$$

la variance :

$$\mathbb{V}_{p(\mathbf{x})} [\mathbf{x}] = \mathbb{E}_{p(\mathbf{x})} \left[\left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \right)^{\text{T}} \left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \right) \right], \quad (1.11)$$

et la matrice de covariance :

$$\mathbb{C}_{p(\mathbf{x})} [\mathbf{x}] = \mathbb{E}_{p(\mathbf{x})} \left[\left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \right) \left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \right)^{\text{T}} \right], \quad (1.12)$$

La variance mesure la moyenne de l'écart quadratique à la moyenne, c'est donc une quantification de l'incertitude dans le sens où, plus la variance est faible, plus les réalisations de \mathbf{x} sont concentrées autour de sa moyenne, donc moins la variable est aléatoire. La matrice de covariance décrit la variance dans chaque direction de l'espace. Notons que les propriétés de la trace induisent :

$$\mathbb{V}_{p(\mathbf{x})} [\mathbf{x}] = \text{tr} \left(\mathbb{C}_{p(\mathbf{x})} [\mathbf{x}] \right). \quad (1.13)$$

Les moments d'ordre supérieurs seront systématiquement ignorés car jugés difficiles à calculer pour une faible utilité pratique. Cette simplification sera à l'origine de la plupart des approximations faites en assimilation de données.

Concernant deux variables aléatoires \mathbf{x} et \mathbf{y} la densité jointe $p(\mathbf{x}, \mathbf{y})$ décrit l'aléatoire dans le couple (\mathbf{x}, \mathbf{y}) . La virgule est à voir comme un "et" logique. Par exemple, la probabilité de " $\mathbf{x} \in E \subseteq \mathbb{R}^n$ et $\mathbf{y} \in F \subseteq \mathbb{R}^d$ " sera :

$$p(\mathbf{x} \in E, \mathbf{y} \in F) = \int_{\mathbf{x} \in E} \int_{\mathbf{y} \in F} p(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} d\mathbf{y}. \quad (1.14)$$

Avec $\mathbf{y} \in \mathbb{R}^d$ l'évènement certain de \mathbf{y} on a :

$$p(\mathbf{x} \in E) = p(\mathbf{x} \in E, \mathbf{y} \in \mathbb{R}^d) = \int_{\mathbf{x} \in E} \int_{\mathbf{y} \in \mathbb{R}^d} p(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} d\mathbf{y}, \quad (1.15)$$

puisque cela est vrai pour toute partie E de \mathbb{R}^n , on obtient la règle de marginalisation :

$$p(\mathbf{x}) = \int p(\mathbf{x}, \mathbf{y}) \, d\mathbf{y}, \quad (1.16)$$

où l'intégrale sans ensemble en indice fait référence à l'intégration sur tout l'espace

d'arrivée de \mathbf{y} . De même, l'espérance de la variable aléatoire $\varphi(\mathbf{x}, \mathbf{y})$ sera notée :

$$\mathbb{E}_{p(\mathbf{x}, \mathbf{y})} [\varphi(\mathbf{x}, \mathbf{y})] = \int \varphi(\mathbf{x}, \mathbf{y}) p(\mathbf{x}, \mathbf{y}) d\mathbf{x}d\mathbf{y}. \quad (1.17)$$

La matrice de covariance de \mathbf{x} et \mathbf{y} sera notée :

$$\mathbb{C}_{p(\mathbf{x}, \mathbf{y})} [\mathbf{x}, \mathbf{y}] = \mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \left[\left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \right) \left(\mathbf{y} - \mathbb{E}_{p(\mathbf{y})} [\mathbf{y}] \right)^{\top} \right]. \quad (1.18)$$

1.1.2 Solution bayésienne

Les probabilités nous permettent aussi de modéliser l'apprentissage ou l'assimilation en décrivant comment la réalisation de l'observation \mathbf{y} influe sur la densité de l'état \mathbf{x} . La densité conditionnelle de \mathbf{x} sachant \mathbf{y} décrit cette influence, elle est définie par :

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{y})}, \quad (1.19)$$

où la densité jointe $p(\mathbf{x}, \mathbf{y})$ est normalisée pour en faire une densité sur \mathbf{x} . Le défaut majeur de cette définition est qu'il ne permet pas de choisir la forme de la densité conditionnelle, ce qui est, comme nous allons le voir, une source de difficultés en pratique¹.

Les occurrences de la variable $\boldsymbol{\varepsilon}$ peuvent ne pas influencer l'aléatoire de \mathbf{x} . Dans ce cas, les variables \mathbf{x} et $\boldsymbol{\varepsilon}$ sont indépendantes. Cela se traduit analytiquement par :

$$p(\mathbf{x}|\boldsymbol{\varepsilon}) = p(\mathbf{x}) \iff \quad (1.20a)$$

$$p(\mathbf{x}, \boldsymbol{\varepsilon}) = p(\mathbf{x}) p(\boldsymbol{\varepsilon}). \quad (1.20b)$$

Une autre caractérisation statistique de cette propriété est que pour toute fonction φ on a :

$$\mathbb{E}_{p(\mathbf{x}|\boldsymbol{\varepsilon})} [\varphi(\mathbf{x})] = \mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})]. \quad (1.21)$$

En considérant les fonctions de la forme $\phi \circ \varphi$, on obtient que les variables $\varphi(\mathbf{x})$ et $\boldsymbol{\varepsilon}$ sont aussi indépendantes. Par symétrie, les variables $\varphi(\mathbf{x})$ et $\psi(\boldsymbol{\varepsilon})$ sont donc indépendantes

1. Une alternative serait d'utiliser un principe variationnel basé sur le minimum d'entropie relative permettant de sélectionner une densité conditionnelle parmi un ensemble de densités candidates (Bishop, 2006).

pour toutes fonctions φ, ψ . Des variables indépendantes sont aussi décorréliées :

$$\mathbb{C}_{p(\mathbf{x}, \boldsymbol{\varepsilon})} [\mathbf{x}, \boldsymbol{\varepsilon}] = \mathbb{E}_{p(\mathbf{x})p(\boldsymbol{\varepsilon})} [\mathbf{x}\boldsymbol{\varepsilon}^T] - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \mathbb{E}_{p(\boldsymbol{\varepsilon})} [\boldsymbol{\varepsilon}]^T, \quad (1.22a)$$

$$= \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \mathbb{E}_{p(\boldsymbol{\varepsilon})} [\boldsymbol{\varepsilon}]^T - \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] \mathbb{E}_{p(\boldsymbol{\varepsilon})} [\boldsymbol{\varepsilon}]^T, \quad (1.22b)$$

$$= \mathbf{0}. \quad (1.22c)$$

La règle de Bayes se déduit de la définition Eq. (1.19) par symétrie de \mathbf{x} et \mathbf{y} (Bernardo and Smith, 2008) :

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})} p(\mathbf{x}), \quad (1.23)$$

elle décrit comment la densité a priori sur l'état $p(\mathbf{x})$ est transformée en la densité a posteriori $p(\mathbf{x}|\mathbf{y})$ par la vraisemblance $p(\mathbf{y}|\mathbf{x})$. La théorie bayésienne des probabilités repose sur cette règle.

Un des objectifs de l'assimilation de données sera donc de calculer cette densité a posteriori. Cependant, l'expression fournie par la règle de Bayes n'a que peu d'utilité en pratique car elle est purement analytique. L'intégration Eq. (1.8) nécessaire au calcul des moments est souvent hors de portée. Il faut alors recourir, comme dans la section suivante, à des approximations.

1.1.3 Méthode par échantillonnage d'importance

La méthode par échantillonnage d'importance (« importance sampling ») consiste à construire un ensemble de M membres $\mathbf{E}^a = [\mathbf{x}^{1,a} \ \dots \ \mathbf{x}^{M,a}] \in \mathbb{R}^{n \times M}$ et un vecteur de M poids positifs $\mathbf{w}^a = [w^{1,a} \ \dots \ w^{M,a}]^T \in \mathbb{R}^M$ pour approcher la densité a posteriori par une somme de distribution de Dirac δ :

$$p(\mathbf{x}|\mathbf{y}) \simeq \sum_{i=1}^M w^{i,a} \delta(\mathbf{x} - \mathbf{x}^{i,a}). \quad (1.24)$$

L'intérêt d'une telle formule est qu'elle permet d'estimer des espérances a posteriori à l'aide de la formule de quadrature (Robert and Casella, 1999) :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\varphi(\mathbf{x})] \simeq \sum_{i=1}^M w^{i,a} \varphi(\mathbf{x}^{i,a}) \quad (1.25a)$$

$$= \varphi(\mathbf{E}^a) \mathbf{w}^a, \quad (1.25b)$$

où l'évaluation de la fonction φ sur l'ensemble \mathbf{E}^a se fait colonne par colonne.

La méthode par échantillonnage d'importance suppose que l'on dispose d'un ensemble

de M membres $\mathbf{E}^b = [\mathbf{x}^{1,b} \ \dots \ \mathbf{x}^{M,b}]$ et poids $\mathbf{w}^b = [w^{1,b} \ \dots \ w^{M,b}]$ permettant d'approcher des espérances a priori via la formule de quadrature :

$$\mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})] \simeq \varphi(\mathbf{E}^b) \mathbf{w}^b. \quad (1.26)$$

Notons que cette expression est insuffisante pour estimer les covariances $\mathbb{C}_{p(\mathbf{x})} [\varphi(\mathbf{x})]$. En effet, $\mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})]$ dans $(\varphi(\mathbf{x}) - \mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})]) (\dots)^T$ (ici, la notation « (\dots) » évite la répétition du facteur de gauche dans le calcul des matrices de covariance) est inconnu avant l'estimation et le remplacer par $\varphi(\mathbf{E}^b) \mathbf{w}^b$ conduit à un estimateur biaisé. On le dé-biaise avec :

$$\mathbb{C}_{p(\mathbf{x})} [\varphi(\mathbf{x})] \simeq \sum_{i=1}^M v^{i,b} (\varphi(\mathbf{x}^{i,b}) - \mathbb{E}_{\mathbf{w}^b} [\varphi(\mathbf{E}^b)]) (\dots)^T, \quad (1.27a)$$

$$v^{i,b} = \frac{w^{i,b}}{\sum_{j=1}^M w^{j,b} (1 - w^{j,b})}, \quad (1.27b)$$

Pour déduire de Eq. (1.26) un ensemble a posteriori, on utilise la règle de Bayes :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\varphi(\mathbf{x})] = \int \varphi(\mathbf{x}) p(\mathbf{x}|\mathbf{y}) d\mathbf{x}, \quad (1.28a)$$

$$\simeq \sum_{i=1}^m \varphi(\mathbf{x}^{i,a}) w^{i,a}, \quad (1.28b)$$

$$= \int \varphi(\mathbf{x}) \frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})} p(\mathbf{x}) d\mathbf{x}, \quad (1.28c)$$

$$\simeq \sum_{i=1}^m \frac{p(\mathbf{y}|\mathbf{x} = \mathbf{x}^{i,b})}{p(\mathbf{y})} w^{i,b} \varphi(\mathbf{x}^{i,b}). \quad (1.28d)$$

Par identification de Eq. (1.28b) avec Eq. (1.28d), une solution simple consiste donc à changer les poids et à garder les mêmes membres :

$$w^{i,a} = \frac{p(\mathbf{y}|\mathbf{x} = \mathbf{x}^{i,b})}{p(\mathbf{y})} w^{i,b}, \quad (1.29a)$$

$$\mathbf{E}^a = \mathbf{E}^b. \quad (1.29b)$$

Pour calculer la constante normalisatrice $p(\mathbf{y})$, on impose que la formule de quadrature

soit exacte pour les fonctions φ constantes :

$$1 = \sum_{i=1}^M w^{i,a}, \Rightarrow \quad (1.30a)$$

$$p(\mathbf{y}) = \sum_{i=1}^M p(\mathbf{y}|\mathbf{x} = \mathbf{x}^{i,b}) w^{i,b}. \quad (1.30b)$$

Pour être efficace, cette méthode demande un très grand nombre de membres lorsque la dimension de \mathbf{x} est grande. Cela en partie parce qu'elle tient compte des moments d'ordre supérieurs à deux. Ce nombre de membres va de pair avec le nombre d'évaluations de l'opérateur \mathcal{H} . En géophysique, cet opérateur contient toute la modélisation physique du système terrestre. On supposera donc que son évaluation est de loin l'opération la plus coûteuse en termes de temps de calcul. Pour être efficace en haute dimension, il faut alors ignorer ces moments d'ordre élevé.

1.1.4 Spécifications gaussiennes a priori

Pour ignorer ces moments d'ordre élevé, les variables d'état \mathbf{x} et d'erreur $\boldsymbol{\varepsilon}$ sont supposées gaussiennes. Un tel choix est justifié depuis deux points de vue dans cette section.

Le point de vue statistique insiste sur les équations entre les variables aléatoires et leurs moments. Si une variable \mathbf{x} est gaussienne de moyenne \mathbf{x}^b et de matrice de covariance \mathbf{B} , on notera :

$$\mathbf{x} \sim \mathcal{N}(\mathbf{x}^b, \mathbf{B}). \quad (1.31)$$

Ici, les propriétés centrales sont que les fonctions affines conservent le caractère gaussien des variables aléatoires :

$$\mathbf{y} = \mathbf{h} + \mathbf{H}\mathbf{x} \Rightarrow \mathbf{y} \sim \mathcal{N}(\mathbf{y}|\mathbf{h} + \mathbf{H}\mathbf{x}^b, \mathbf{H}\mathbf{B}\mathbf{H}^T), \quad (1.32)$$

et que deux variables aléatoires décorréées sont indépendantes :

$$\mathbb{C}_{p(\mathbf{x},\boldsymbol{\varepsilon})}[\mathbf{x}, \boldsymbol{\varepsilon}] = \mathbf{0} \Rightarrow \mathbf{x}, \boldsymbol{\varepsilon} \text{ sont indépendantes.} \quad (1.33)$$

Ici, l'utilisation de variables aléatoires gaussiennes est justifiée par le théorème central limite. Il stipule que la distribution d'une variable aléatoire qui est la somme d'un grand nombre de causes indépendantes est gaussienne. Ce théorème explique donc la présence naturelle de la distribution gaussienne dans des situations complexes.

Le point de vue analytique insiste sur les propriétés de la densité $p(\mathbf{x})$. Si $p(\mathbf{x})$ est

une densité gaussienne, on notera :

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{x}^b, \mathbf{B}). \quad (1.34)$$

Ici, les propriétés centrales sont que la fonction de coût $G(\mathbf{x})$ associée à $p(\mathbf{x})$:

$$G(\mathbf{x}) \propto -\ln p(\mathbf{x}), \quad (1.35)$$

ou le symbole \propto signifie « égal à une constante additive près » lorsqu'il est appliqué à des fonction de coût, est quadratique définie-positive :

$$G(\mathbf{x}) \propto \frac{1}{2} \|\mathbf{x} - \mathbf{x}^*\|_{\nabla^2 G}^2. \quad (1.36)$$

La notation :

$$\frac{1}{2} \|\mathbf{x} - \mathbf{x}^*\|_{\nabla^2 G}^2 = (\mathbf{x} - \mathbf{x}^*)^T \nabla_{\mathbf{x}}^2 G (\mathbf{x} - \mathbf{x}^*), \quad (1.37)$$

est le carré de la norme euclidienne engendrée par la matrice hessienne $\nabla_{\mathbf{x}}^2 G$ de G qui est symétrique, définie-positive et constante. Le point \mathbf{x}^* est, dans ce cas, l'unique minimiseur global de la fonction de coût. Il est caractérisé par l'équation de gradient nul :

$$\nabla_{\mathbf{x}} G(\mathbf{x}^*) = \mathbf{0}_n. \quad (1.38)$$

L'Eq. (1.36) est donc le développement de Taylor d'ordre deux de G . L'utilisation de la densité gaussienne est ici justifiée par la maximisation de l'entropie. En effet, parmi toutes les densités de moyenne et de matrice de covariance fixées, la densité gaussienne est l'unique maximiseur de l'entropie (Cf annexe A.1). Or, l'entropie est une mesure de l'incertitude, de l'information ou de la quantité de connaissances dont on dispose sur une variable aléatoire. Ainsi, si toute notre connaissance de la variable aléatoire \mathbf{x} se résume à sa moyenne et à sa matrice de covariance, on supposera sa densité gaussienne car tout autre choix de densité introduirait une forme d'information artificielle.

Un lien entre ces deux points de vue se fait en appliquant Eq. (1.9) à $\varphi(\mathbf{x}) = 1$ et $\varphi(\mathbf{x}) = \nabla_{\mathbf{x}} G(\mathbf{x})$. On obtient :

$$\mathbb{E}_{p(\mathbf{x})} [\nabla_{\mathbf{x}} G(\mathbf{x})] = \mathbf{0}_n, \quad (1.39a)$$

$$\mathbb{C}_{p(\mathbf{x})} [\nabla_{\mathbf{x}} G(\mathbf{x})] = \mathbb{E}_{p(\mathbf{x})} [\nabla_{\mathbf{x}}^2 G(\mathbf{x})]. \quad (1.39b)$$

Dans notre cas où $G(\mathbf{x})$ est une fonction quadratique, son gradient est une fonction

affine, sa matrice hessienne une fonction constante d'où :

$$\nabla_{\mathbf{x}} G(\mathbf{x}^b) = \mathbf{0}_n, \quad (1.40a)$$

$$\left(\nabla_{\mathbf{x}}^2 G\right) \mathbf{B} \left(\nabla_{\mathbf{x}}^2 G\right) = \nabla_{\mathbf{x}}^2 G. \quad (1.40b)$$

Ce qui nous permet d'affirmer par unicité que la moyenne et la matrice de covariance d'une variable aléatoire gaussienne sont respectivement le minimiseur et l'inverse de la hessienne de sa fonction de coût :

$$\mathbb{E}_{p(\mathbf{x})}[\mathbf{x}] = \mathbf{x}^b = \mathbf{x}^*, \quad (1.41a)$$

$$\mathbb{C}_{p(\mathbf{x})}[\mathbf{x}] = \mathbf{B}^{-1} = \left(\nabla_{\mathbf{x}}^2 G\right)^{-1}. \quad (1.41b)$$

Pour terminer de spécifier l'aléatoire du système d'assimilation, il faut aussi attribuer une densité à l'erreur d'observation $\boldsymbol{\varepsilon}$. Si notre connaissance de l'aléatoire de l'erreur d'observation $\boldsymbol{\varepsilon}$ se résume à sa moyenne \mathbf{e}^r et sa matrice de covariance \mathbf{R} , le même argument de maximisation de l'entropie demande de supposer :

$$p(\boldsymbol{\varepsilon}) = \mathcal{N}(\boldsymbol{\varepsilon} | \mathbf{e}^r, \mathbf{R}). \quad (1.42)$$

Cependant, si $\mathbf{e}^r \neq \mathbf{0}_d$ l'opérateur d'observation est biaisé. Quitte à le débiaiser, on peut supposer $\mathbf{e}^r = \mathbf{0}_d$. Plus généralement, quitte à ajouter à \mathcal{H} une fonction affine de \mathbf{x} , on peut supposer $\boldsymbol{\varepsilon}$ de moyenne nulle et décorrélée de \mathbf{x} . Dans ce cas, $\boldsymbol{\varepsilon}$ et \mathbf{x} sont indépendantes et \mathcal{H} est le meilleur estimateur de \mathbf{y} fonction de \mathbf{x} d'après l'annexe A.3.2. Ces hypothèses statistiques nous donnent l'expression analytique de la vraisemblance par changement de variable :

$$p(\mathbf{y} | \mathbf{x}) \propto e^{-G(\mathbf{y} | \mathbf{x})}, \quad (1.43a)$$

$$G(\mathbf{y} | \mathbf{x}) \propto \frac{1}{2} \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|_{\mathbf{R}^{-1}}^2, \quad (1.43b)$$

Notons que la matrice hessienne en \mathbf{y} de la fonction de coût $\nabla_{\mathbf{y}}^2 G(\mathbf{y} | \mathbf{x}) = \mathbf{R}^{-1}$ a la particularité d'être une fonction constante de \mathbf{x} et \mathbf{y} .

1.1.5 Échantillonner une densité

Dans la méthode par échantillonnage d'importance de la section 1.1.3, on supposait disposer d'un ensemble a priori \mathbf{E}^b permettant d'approcher $p(\mathbf{x})$. Plusieurs méthodes sont développées ici pour engendrer un tel ensemble.

Lorsque $p(\mathbf{x})$ est gaussienne, la méthode stochastique utilise un générateur de nombres aléatoires permettant d'obtenir un ensemble $\mathbf{U}_M = (\mathbf{u}^1, \dots, \mathbf{u}^M) \in \mathbb{R}^{n \times M}$ de membres indépendants et distribués selon $\mathcal{N}(\mathbf{u} | \mathbf{0}_n, \mathbf{I}_n)$ avec $\mathbf{0}_n$ le vecteur nul de \mathbb{R}^n et \mathbf{I}_n la matrice identité de $\mathbb{R}^{n \times n}$. L'ensemble stochastique \mathbf{E}^b est alors :

$$\mathbf{E}^b = \mathbf{x}^b \mathbf{1}_M^T + \mathbf{X}^b \mathbf{U}_M = \begin{bmatrix} \mathbf{x}^b & \mathbf{X}^b \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \mathbf{U}_M \end{bmatrix}^T, \quad (1.44)$$

où le vecteur $\mathbf{1}_M \in \mathbb{R}^M$ ne contient que des 1 et la matrice $\mathbf{X}^b \in \mathbb{R}^{n \times n}$ est un facteur de \mathbf{B} tel que

$$(\mathbf{X}^b) (\mathbf{X}^b)^T = \mathbf{B}. \quad (1.45)$$

Un tel facteur est accessible via une décomposition de Cholesky ou une diagonalisation (Stoer and Bulirsch, 1982) et les poids sont choisis égaux à $\frac{1}{M}$. Grâce à ce type d'échantillonnage, l'approximation Eq. (1.25a) devient de plus en plus précise à mesure que le nombre de particules M augmente et cela quel que soit φ (dans L^2) :

$$\mathbb{E}_{p(\mathbf{x}^1, \dots, \mathbf{x}^M, \mathbf{b})} \left[\left\| \mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})] - \varphi(\mathbf{E}^b) \frac{\mathbf{1}_M}{M} \right\|^2 \right] \quad (1.46a)$$

$$= \frac{1}{M^2} \mathbb{E}_{p(\mathbf{x}^1, \dots, \mathbf{x}^M, \mathbf{b})} \left[\left\| \sum_{i=1}^M (\mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})] - \varphi(\mathbf{x}^{i, \mathbf{b}})) \right\|^2 \right], \quad (1.46b)$$

$$= \frac{1}{M^2} \sum_{i=1}^M \mathbb{E}_{p(\mathbf{x}^{i, \mathbf{b}})} \left[\left\| \mathbb{E}_{p(\mathbf{x})} [\varphi(\mathbf{x})] - \varphi(\mathbf{x}^{i, \mathbf{b}}) \right\|^2 \right], \quad (1.46c)$$

$$= \frac{1}{M} \mathbb{V}_{p(\mathbf{x})} [\varphi(\mathbf{x})], \quad (1.46d)$$

où l'on a utilisé l'indépendance mutuelle des $\mathbf{x}^{i, \mathbf{b}}$ pour intervertir somme et espérance du carré de la norme.

On peut cependant reprocher à cette méthode de nécessiter un grand nombre de particules pour pouvoir intégrer des fonctions φ simplement affines. Pour contourner le problème, on peut engendrer un ensemble semi-stochastique $\mathbf{E}^{b'}$ tel que sa moyenne empirique soit exacte. Pour cela, il suffit de recentrer l'ensemble précédent à l'aide de :

$$\mathbf{E}^{b'} = \mathbf{x}^b \mathbf{1}_M^T + \mathbf{E}^{b'} \left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right). \quad (1.47)$$

Une telle définition pour $\mathbf{E}^{b'}$ donne en effet une quadrature exacte pour les fonctions φ

affines :

$$\varphi\left(\mathbf{E}^{b'}\right) \frac{\mathbf{1}_M}{M} = \varphi\left(\mathbf{E}^{b'} \frac{\mathbf{1}_M}{M}\right), \quad (1.48a)$$

$$= \varphi\left(\mathbf{x}^b\right), \quad (1.48b)$$

$$= \varphi\left(\mathbb{E}_{p(\mathbf{x})}[\mathbf{x}]\right), \quad (1.48c)$$

$$= \mathbb{E}_{p(\mathbf{x})}[\varphi(\mathbf{x})]. \quad (1.48d)$$

En continuant sur cette voie, on peut vouloir engendrer un ensemble déterministe $\mathbf{E}^{b''} = (\mathbf{x}^{b'',1}, \dots, \mathbf{x}^{b'',M}) \in \mathbb{R}^{n \times M}$ tel que sa moyenne et sa matrice de covariance empiriques soient exactes (ou les plus précises possible). Si ces moments empiriques sont calculés avec 1.26 et 1.27a en utilisant des poids constants, l'ensemble $\mathbf{E}^{b''}$ doit alors vérifier :

$$\mathbf{E}^{b''} \frac{\mathbf{1}_M}{M} = \mathbf{x}^b, \quad (1.49a)$$

$$\mathbf{E}^{b''} \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{\sqrt{m}} \right) \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{\sqrt{m}} \right) (\mathbf{E}^{b''})^T = \mathbf{B}, \quad (1.49b)$$

avec $m = M - 1$. Soit $\mathbf{X}^{b''} \in \mathbb{R}^{n \times m}$ une certaine matrice, posons :

$$\mathbf{E}^{b''} = \begin{bmatrix} \mathbf{x}^b & \mathbf{X}^{b''} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T, \quad (1.50)$$

avec $\mathbf{U}_M \in \mathbb{R}^{M \times m}$ telle que la matrice $\begin{bmatrix} \frac{\mathbf{1}_M}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix} \in \mathbb{R}^{M \times M}$ soit orthogonale. L'inverse matricielle de $\begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$ étant $\begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} \end{bmatrix}$, on obtient ainsi :

$$\begin{bmatrix} \mathbf{x}^b & \mathbf{X}^{b''} \end{bmatrix} = \mathbf{E}^{b''} \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} \end{bmatrix}. \quad (1.51)$$

La moyenne empirique $\mathbf{E}^{b''} \frac{\mathbf{1}_M}{M}$ de $\mathbf{E}^{b''}$ est donc bien \mathbf{x}^b . En ce qui concerne $\mathbf{X}^{b''} = \mathbf{E}^{b''} \frac{\mathbf{U}_M}{\sqrt{m}}$, notons que l'orthogonalité de $\begin{bmatrix} \frac{\mathbf{1}_M}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}$ implique :

$$\begin{bmatrix} \frac{\mathbf{1}_M}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix} \begin{bmatrix} \frac{\mathbf{1}_M}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}^T = \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} + \mathbf{U}_M \mathbf{U}_M^T, \quad (1.52a)$$

$$= \mathbf{I}_M. \quad (1.52b)$$

donc :

$$\mathbf{U}_M \mathbf{U}_M^T = \mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}, \quad (1.53a)$$

$$= \left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right) \left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right)^T, \quad (1.53b)$$

puisque la matrice $\left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right)$ est symétrique et idempotente. Ainsi, avec :

$$\mathbf{A}^{b''} = \mathbf{E}^{b''} \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{\sqrt{m}} \right), \quad (1.54)$$

la matrice d'anomalie normalisée de $\mathbf{E}^{b''}$, on a :

$$\left(\mathbf{A}^{b''} \right) \left(\mathbf{A}^{b''} \right)^T = \left(\mathbf{E}^{b''} \right) \frac{\mathbf{U}_M}{\sqrt{m}} \frac{\mathbf{U}_M^T}{\sqrt{m}} \left(\mathbf{E}^{b''} \right)^T, \quad (1.55a)$$

$$= \left(\mathbf{X}^{b''} \right) \left(\mathbf{X}^{b''} \right)^T. \quad (1.55b)$$

La covariance empirique de $\mathbf{E}^{b''}$ est donc $\left(\mathbf{X}^{b''} \right) \left(\mathbf{X}^{b''} \right)^T$. Dans la suite, une telle matrice $\mathbf{X}^{b''} = \mathbf{E}^{b''} \frac{\mathbf{U}_M}{\sqrt{m}} \in \mathbb{R}^{M \times m}$ sera nommée : « une matrice de déviation de $\mathbf{E}^{b''}$ » par opposition à sa matrice d'anomalie. Pour que $\mathbf{E}^{b''}$ soit solution de Eq. (1.49), il suffit donc d'avoir :

$$\left(\mathbf{X}^{b''} \right) \left(\mathbf{X}^{b''} \right)^T = \mathbf{B}. \quad (1.56)$$

Cependant, un tel facteur $\mathbf{X}^{b''} \in \mathbb{R}^{M \times m}$ n'existe pas toujours lorsque $m < n$ puisque \mathbf{B} est de rang n et $\left(\mathbf{X}^{b''} \right) \left(\mathbf{X}^{b''} \right)^T$ de rang au plus m . Dans ce cas, le théorème d'Eckart-Young (Björck, 1996) montre que la meilleure approximation de rang au plus m en norme de Frobenius s'obtient à l'aide d'une décomposition spectrale de la matrice symétrique définie-positive \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} \mathbf{V}_0 & \mathbf{V}_1 \end{bmatrix} \begin{bmatrix} \mathbf{D}_0^2 & \\ & \mathbf{D}_1^2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_0 & \mathbf{V}_1 \end{bmatrix}^T, \quad (1.57a)$$

$$\mathbf{X}^{b''} = \mathbf{V}_0 \mathbf{D}_0. \quad (1.57b)$$

où la matrice diagonale $\mathbf{D}_0^2 \in \mathbb{R}^{m \times m}$ contient les m plus grandes valeurs propres de \mathbf{B} et la matrice diagonale $\mathbf{D}_1^2 \in \mathbb{R}^{(n-m) \times (n-m)}$ le reste. Les matrices $\mathbf{V}_0 \in \mathbb{R}^{n \times m}$, $\mathbf{V}_1 \in \mathbb{R}^{n \times (n-m)}$ contiennent les vecteurs propres associés tels que $\begin{bmatrix} \mathbf{V}_0 & \mathbf{V}_1 \end{bmatrix} \in \mathbb{R}^{n \times n}$ soit orthogonale.

Les ensembles ont donc ici un double intérêt. D'abord ils permettent, à l'aide des formules de quadrature, d'estimer des espérances. Ensuite, ils permettent d'encoder sur peu de membres la donnée de la moyenne et de la matrice de covariance d'une densité. Ceci constitue un réel avantage en géophysique où les matrices de covariance de variables à plusieurs millions de dimensions sont difficilement manipulables en mémoire.

1.1.6 Solution dans le cadre linéaire et gaussien

A l'aide des hypothèses de la section 1.1.4 et de la règle de Bayes Eq. (1.23) une expression de la densité jointe est :

$$p(\mathbf{x}, \mathbf{y}) \propto e^{-G(\mathbf{x}, \mathbf{y})}, \quad (1.58)$$

$$G(\mathbf{x}, \mathbf{y}) \propto \frac{1}{2} \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathbf{x} - \mathbf{x}^b\|_{\mathbf{B}^{-1}}^2. \quad (1.59)$$

Nous montrons dans cette section l'équivalence entre le caractère affine de \mathcal{H} et le caractère gaussien de (\mathbf{x}, \mathbf{y}) . Dans chaque implication, nous donnerons une expression de la moyenne et de la covariance a posteriori.

La première implication utilise une propriété analytique de \mathcal{H} , nous utiliserons donc le point de vue analytique de la section 1.1.4. Si \mathcal{H} est une fonction affine de \mathbf{x} :

$$\mathcal{H}(\mathbf{x}) = \mathbf{h} + \mathbf{H}(\mathbf{x} - \mathbf{x}^b), \quad (1.60)$$

alors $G(\mathbf{x}, \mathbf{y})$ est une fonction quadratique, définie positive de \mathbf{x} et \mathbf{y} donc $p(\mathbf{x}, \mathbf{y})$ est la gaussienne de moments :

$$\mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{pmatrix} \mathbf{x}^b \\ \mathbf{h} \end{pmatrix}, \quad (1.61a)$$

$$\mathbb{C}_{p(\mathbf{x}, \mathbf{y})} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{pmatrix} \mathbf{B} & \mathbf{B}\mathbf{H}^T \\ \mathbf{H}\mathbf{B} & \mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R} \end{pmatrix}. \quad (1.61b)$$

Puisque $p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{y})}$ d'après Eq. (1.19), la fonction de coût a posteriori $G(\mathbf{x}|\mathbf{y})$ est égale à $G(\mathbf{x}, \mathbf{y})$ à une constante de la variable \mathbf{x} près :

$$G(\mathbf{x}|\mathbf{y}) \propto_{\mathbf{x}} G(\mathbf{x}, \mathbf{y}). \quad (1.62)$$

C'est donc une fonction quadratique, définie-positive de \mathbf{x} . Ainsi, la densité a posteriori

$p(\mathbf{x}|\mathbf{y})$ est gaussienne. En calculant les dérivées de $G(\mathbf{x}|\mathbf{y})$ selon \mathbf{x} et en résolvant l'équation de gradient nulle Eq. (1.38) par un développement de Taylor :

$$\nabla_{\mathbf{x}}G(\mathbf{x}^*|\mathbf{y}) = \mathbf{0}_n \Leftrightarrow, \quad (1.63a)$$

$$\nabla_{\mathbf{x}}G(\mathbf{x}|\mathbf{y}) + \nabla_{\mathbf{x}}^2G \times (\mathbf{x} - \mathbf{x}^*) = \mathbf{0}_n, \quad (1.63b)$$

on obtient, par évaluation en $\mathbf{x} = \mathbf{x}^b$, le minimiseur a posteriori (ou MAP) et la matrice hessienne de la fonction de coût :

$$\mathbf{x}^* = \mathbf{x}^b + \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)^{-1} \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}), \quad (1.64a)$$

$$\nabla_{\mathbf{x}}^2G = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}. \quad (1.64b)$$

Ce qui permet d'obtenir les moments a posteriori à l'aide d'Eq. (1.41a,1.41b) :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] = \mathbf{x}^b + \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)^{-1} \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}), \quad (1.65a)$$

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)^{-1}. \quad (1.65b)$$

Pour la seconde implication, nous utiliserons le point de vue statistique de la section 1.1.4. Si le couple (\mathbf{x}, \mathbf{y}) est gaussien, il existe une fonction affine $\tilde{\mathcal{H}}$ et une erreur d'observation $\tilde{\boldsymbol{\varepsilon}}$ de moyenne nulle et indépendante de \mathbf{x} telles que $\mathbf{y} = \tilde{\mathcal{H}}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}}$. En effet, définissons $\tilde{\mathcal{H}}(\mathbf{x}) = \mathbf{h} + \mathbf{H}(\mathbf{x} - \mathbf{x}^b)$ et $\tilde{\boldsymbol{\varepsilon}}$ avec :

$$\mathbf{h} = \mathbb{E}_{p(\mathbf{y})}[\mathbf{y}], \quad (1.66a)$$

$$\mathbf{H} = \mathbb{C}_{p(\mathbf{x},\mathbf{y})}[\mathbf{y}, \mathbf{x}] \left(\mathbb{C}_{p(\mathbf{x})}[\mathbf{x}]\right)^{-1}, \quad (1.66b)$$

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{y} - \tilde{\mathcal{H}}(\mathbf{x}). \quad (1.66c)$$

Le couple $(\mathbf{x}, \tilde{\boldsymbol{\varepsilon}})$ est gaussien en tant que fonction affine du couple (\mathbf{x}, \mathbf{y}) . À l'aide des propriétés de l'espérance, on montre que $\tilde{\boldsymbol{\varepsilon}}$ est de moyenne nulle :

$$\mathbb{E}_{p(\tilde{\boldsymbol{\varepsilon}})}[\tilde{\boldsymbol{\varepsilon}}] = \mathbb{E}_{p(\mathbf{x},\mathbf{y})}[\mathbf{y} - \tilde{\mathcal{H}}(\mathbf{x})], \quad (1.67a)$$

$$= \mathbb{E}_{p(\mathbf{y})}[\mathbf{y}] - \mathbf{h}, \quad (1.67b)$$

$$= \mathbf{0}_d, \quad (1.67c)$$

et que le couple $(\mathbf{x}, \tilde{\boldsymbol{\varepsilon}})$ est décorrélé :

$$\mathbb{C}_{p(\mathbf{x}, \mathbf{y})} [\mathbf{y}, \mathbf{x}] = \mathbb{C}_{p(\mathbf{x}, \mathbf{y})} [\tilde{\mathcal{H}}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}}, \mathbf{x}], \quad (1.68)$$

$$= \mathbf{H} \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}] + \mathbb{C}_{p(\tilde{\boldsymbol{\varepsilon}}, \mathbf{x})} [\tilde{\boldsymbol{\varepsilon}}, \mathbf{x}] \Rightarrow \quad (1.69)$$

$$\mathbb{C}_{p(\tilde{\boldsymbol{\varepsilon}}, \mathbf{x})} [\tilde{\boldsymbol{\varepsilon}}, \mathbf{x}] = \mathbf{0}_{d \times n}, \quad (1.70)$$

donc indépendant d'après Eq. (1.33). En appliquant l'espérance conditionnelle à l'équation $\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon} = \tilde{\mathcal{H}}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}}$ on trouve :

$$\mathbb{E}_{p(\mathbf{y}|\mathbf{x})} [\mathbf{y}] = \mathcal{H}(\mathbf{x}) = \tilde{\mathcal{H}}(\mathbf{x}), \quad (1.71)$$

donc \mathcal{H} est une fonction affine. Par symétrie des rôles joués par \mathbf{x} et \mathbf{y} dans $p(\mathbf{x}, \mathbf{y})$, la fonction affine $\mathcal{K}(\mathbf{y}) = \mathbf{k} + \mathbf{K}(\mathbf{y} - \mathbb{E}_{p(\mathbf{y})}[\mathbf{y}])$ et l'erreur d'estimation $\boldsymbol{\eta}$ définies par :

$$\mathbf{k} = \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}], \quad (1.72a)$$

$$\mathbf{K} = \mathbb{C}_{p(\mathbf{x}, \mathbf{y})} [\mathbf{x}, \mathbf{y}] \mathbb{C}_{p(\mathbf{y})} [\mathbf{y}]^{-1}, \quad (1.72b)$$

$$\boldsymbol{\eta} = \mathbf{x} - \mathcal{K}(\mathbf{y}), \quad (1.72c)$$

sont telles que $\boldsymbol{\eta}$ est gaussienne de moyenne nulle et décorrélée de \mathbf{y} . Donc $\boldsymbol{\eta}$ et \mathbf{y} sont indépendantes et $\mathcal{K}(\mathbf{y}) = \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}]$. La matrice \mathbf{K} est aussi appelée gain de Kalman. Cela permet, d'après l'annexe A.3.2, de montrer que \mathcal{K} est le meilleur estimateur de \mathbf{x} fonction de \mathbf{y} en erreur quadratique moyenne (ou MMSE). Puisque \mathcal{K} est une fonction affine de \mathbf{y} , c'est aussi le meilleur estimateur linéaire non biaisé de \mathbf{x} (ou BLUE, e.g. Puntanen and Styan, 2011). La décomposition $\mathbf{x} = \mathcal{K}(\mathbf{y}) + \boldsymbol{\eta}$ permet d'exprimer la moyenne et la covariance a posteriori à l'aide des propriétés de l'espérance conditionnelle,

ces expressions sont ensuite précisées avec Eq. (1.72,1.61) :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}] = \mathcal{K}(\mathbf{y}) + \mathbb{E}_{p(\boldsymbol{\eta}|\mathbf{y})} [\boldsymbol{\eta}], \quad (1.73a)$$

$$= \mathcal{K}(\mathbf{y}) + \mathbb{E}_{p(\boldsymbol{\eta})} [\boldsymbol{\eta}], \quad (1.73b)$$

$$= \mathcal{K}(\mathbf{y}), \quad (1.73c)$$

$$= \mathbb{E}_{p(\mathbf{x})} [\mathbf{x}] + \mathbb{C}_{p(\mathbf{x},\mathbf{y})} [\mathbf{x}, \mathbf{y}] \mathbb{C}_{p(\mathbf{y})} [\mathbf{y}]^{-1} (\mathbf{y} - \mathbb{E}_{p(\mathbf{y})} [\mathbf{y}]), \quad (1.73d)$$

$$= \mathbf{x}^b + \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{h}), \quad (1.73e)$$

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}] = \mathbb{C}_{p(\boldsymbol{\eta}|\mathbf{y})} [\boldsymbol{\eta}], \quad (1.73f)$$

$$= \mathbb{C}_{p(\boldsymbol{\eta})} [\boldsymbol{\eta}], \quad (1.73g)$$

$$= \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}] - \mathbf{K}\mathbb{C}_{p(\mathbf{y})} [\mathbf{y}] \mathbf{K}^T, \quad (1.73h)$$

$$= \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}] - \mathbb{C}_{p(\mathbf{x},\mathbf{y})} [\mathbf{x}, \mathbf{y}] \mathbb{C}_{p(\mathbf{y})} [\mathbf{y}]^{-1} \mathbb{C}_{p(\mathbf{x},\mathbf{y})} [\mathbf{y}, \mathbf{x}], \quad (1.73i)$$

$$= \mathbf{B} - \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{B}. \quad (1.73j)$$

L'égalité entre Eq. (1.65a,1.65b) et Eq. (1.73e,1.73j) démontre les formules de Sherman-Morrison-Woodbury (Asch et al., 2017) :

$$\left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1}\right)^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}, \quad (1.74a)$$

$$\left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1}\right)^{-1} = \mathbf{B} - \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{B}. \quad (1.74b)$$

Lorsque $d < n$ les expressions des membres de droite sont plus faciles à calculer puisque la matrice à inverser est de taille $d \times d$ plutôt que $n \times n$. Ces expressions de la matrice de covariance a posteriori permettent aussi de remarquer que,

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})} \leq \mathbb{C}_{p(\mathbf{x})}, \quad (1.75)$$

avec \leq l'ordre partiel des matrices symétriques, semi-définies-positives. Ce qui signifie que l'assimilation de \mathbf{y} a réduit la variance et donc l'incertitude sur \mathbf{x} dans toutes les directions de l'espace. La figure 1.1 illustre cette diminution de l'incertitude

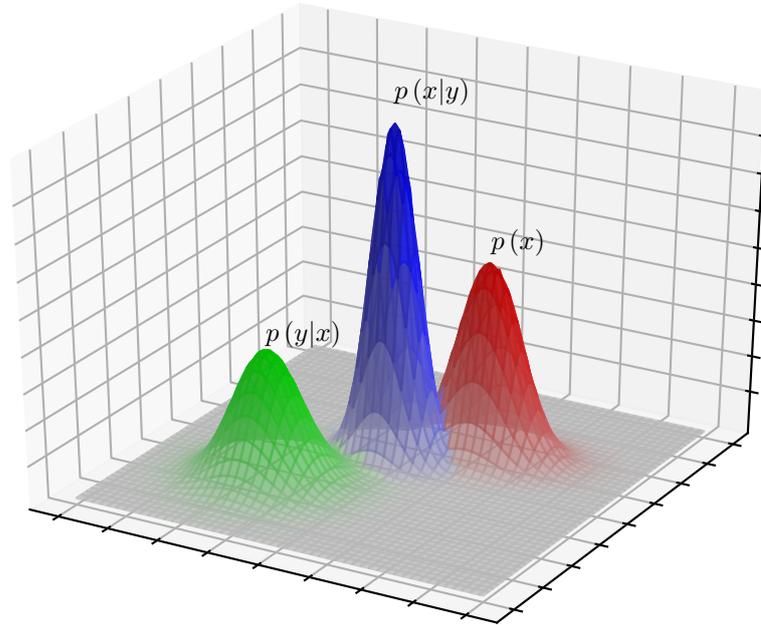


FIGURE 1.1 – Transformation de la densité a priori $p(\mathbf{x})$ en densité a posteriori $p(\mathbf{x}|\mathbf{y})$ à l'aide de la vraisemblance $p(\mathbf{y}|\mathbf{x})$.

Dans le cadre linéaire et gaussien, la moyenne et la matrice de covariance a posteriori sont importantes du point de vue analytique parce qu'elles caractérisent la densité a posteriori. D'un point de vue statistique, la moyenne a posteriori est l'estimateur de \mathbf{x} fonction affine de \mathbf{y} dont l'erreur est indépendante de \mathbf{y} . Voyons à présent ce qu'il en est dans un cadre moins linéaire et gaussien.

1.1.7 Méthode dans un cadre moins linéaire et gaussien

Lorsque \mathcal{H} n'est plus affine et que le couple (\mathbf{x}, \mathbf{y}) n'est plus gaussien, la densité a posteriori n'est plus caractérisée par sa moyenne et sa matrice de covariance. Il n'existe plus d'estimateur de \mathbf{x} fonction affine de \mathbf{y} dont l'erreur est indépendante de \mathbf{y} . Ces moments sont néanmoins toujours importants parce qu'ils engendrent la meilleure approximation

gaussienne de la densité a posteriori :

$$p(\mathbf{x}|\mathbf{y}) \simeq \mathcal{N}\left(\mathbf{x}|\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}], \mathbb{C}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}]\right), \quad (1.76)$$

au sens de l'entropie relative (annexe A.2). En particulier, la moyenne a posteriori $\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}]$ est le meilleur estimateur de \mathbf{x} fonction de \mathbf{y} au sens de l'erreur quadratique moyenne. Malheureusement, l'intégration nécessaire au calcul de ces moments a posteriori ne peut être conduite algébriquement. En revanche, on peut revenir au cas de la section 1.1.6 à l'aide d'une approximation gaussienne de la densité jointe $p(\mathbf{x}, \mathbf{y})$. Si le critère d'approximation de densité est l'entropie relative, l'annexe A.2 montre que la meilleure approximation gaussienne de la densité jointe est :

$$p(\mathbf{x}, \mathbf{y}) \simeq q(\mathbf{x}, \mathbf{y}), \quad (1.77a)$$

$$q(\mathbf{x}, \mathbf{y}) = \mathcal{N}\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \middle| \mathbb{E} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \mathbb{C} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}\right). \quad (1.77b)$$

Dans le cadre de cette approximation et d'après Eq. (1.73d,1.73i), les moments a posteriori sont alors approchés par :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] \simeq \mathcal{K}(\mathbf{y}), \quad (1.78a)$$

$$\mathcal{K}(\mathbf{y}) = \mathbb{E}_{p(\mathbf{x})}[\mathbf{x}] + \mathbb{C}_{p(\mathbf{x},\mathbf{y})}[\mathbf{x}, \mathbf{y}] \mathbb{C}_{p(\mathbf{y})}[\mathbf{y}]^{-1} (\mathbf{y} - \mathbb{E}_{p(\mathbf{y})}[\mathbf{y}]), \quad (1.78b)$$

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] \simeq \mathbb{C}_{p(\tilde{\boldsymbol{\eta}})}[\boldsymbol{\eta}], \quad (1.78c)$$

$$\mathbb{C}_{p(\tilde{\boldsymbol{\eta}})}[\boldsymbol{\eta}] = \mathbb{C}_{p(\mathbf{x})}[\mathbf{x}] - \mathbb{C}_{p(\mathbf{x},\mathbf{y})}[\mathbf{x}, \mathbf{y}] \mathbb{C}_{p(\mathbf{y})}[\mathbf{y}]^{-1} \mathbb{C}_{p(\mathbf{x},\mathbf{y})}[\mathbf{y}, \mathbf{x}]. \quad (1.78d)$$

avec $\boldsymbol{\eta} = \mathbf{x} - \mathcal{K}(\mathbf{y})$ l'erreur d'estimation non gaussienne qui est de moyenne nulle et seulement décorrélée de \mathbf{y} . D'après l'annexe A.3, la fonction affine \mathcal{K} est alors le meilleur estimateur de \mathbf{x} fonction affine de \mathbf{y} en erreur quadratique moyenne (BLUE). Grâce à l'indépendance de \mathbf{x} et $\boldsymbol{\varepsilon}$, les moments de la densité jointe Eq. (1.77b) nécessitent seulement une intégration par rapport à la densité a priori :

$$\mathbb{E} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{pmatrix} \mathbb{E}_{p(\mathbf{x})}[\mathbf{x}] \\ \mathbb{E}_{p(\mathbf{x})}[\mathcal{H}(\mathbf{x})] \end{pmatrix}, \quad (1.79a)$$

$$\mathbb{C} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{pmatrix} \mathbb{C}_{p(\mathbf{x})}[\mathbf{x}] & \mathbb{C}_{p(\mathbf{x})}[\mathbf{x}, \mathcal{H}(\mathbf{x})] \\ \mathbb{C}_{p(\mathbf{x})}[\mathcal{H}(\mathbf{x}), \mathbf{x}] & \mathbb{C}_{p(\mathbf{x})}[\mathcal{H}(\mathbf{x})] + \mathbf{R} \end{pmatrix}. \quad (1.79b)$$

Notons que, contrairement au cas linéaire :

$$\mathbb{C}_{p(\mathbf{x})} [\mathcal{H}(\mathbf{x})] \neq \mathbb{C}_{p(\mathbf{x})} [\mathcal{H}(\mathbf{x}), \mathbf{x}] \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}]^{-1} \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}, \mathcal{H}(\mathbf{x})]. \quad (1.80)$$

D'après la section 1.1.5, on sait engendrer un échantillon \mathbf{E}^b de $p(\mathbf{x})$. On peut alors estimer les moments Eq. (1.79) à l'aide des Eq. (1.26,1.27a) :

$$\mathbb{E}_{p(\mathbf{x})} \begin{bmatrix} \mathbf{x} \\ \mathcal{H}(\mathbf{x}) \end{bmatrix} \simeq \begin{pmatrix} \mathbf{E}^b \\ \mathcal{H}(\mathbf{E}^b) \end{pmatrix} \frac{\mathbf{1}_M}{M}, \quad (1.81a)$$

$$\mathbb{C}_{p(\mathbf{x})} \begin{bmatrix} \mathbf{x} \\ \mathcal{H}(\mathbf{x}) \end{bmatrix} \simeq \begin{pmatrix} \mathbf{E}^b \\ \mathcal{H}(\mathbf{E}^b) \end{pmatrix} \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{M-1} \right) \begin{pmatrix} \mathbf{E}^b \\ \mathcal{H}(\mathbf{E}^b) \end{pmatrix}^T. \quad (1.81b)$$

Puisqu'il y a équivalence entre le caractère gaussien de (\mathbf{x}, \mathbf{y}) et le caractère affine de \mathcal{H} , on peut se demander si il existe une approximation affine $\tilde{\mathcal{H}}(\mathbf{x}) = \tilde{\mathbf{h}} + \tilde{\mathbf{H}}(\mathbf{x} - \mathbf{x}^b)$ de \mathcal{H} qui correspondrait à l'approximation gaussienne Eq. (1.77) de $p(\mathbf{x}, \mathbf{y})$. Une telle approximation affine de l'opérateur d'observation n'est cependant pas suffisante pour engendrer une approximation gaussienne du couple (\mathbf{x}, \mathbf{y}) . On lui associe donc une approximation gaussienne de l'erreur d'estimation $\tilde{\boldsymbol{\varepsilon}} = \mathbf{y} - \tilde{\mathcal{H}}(\mathbf{x})$:

$$p(\tilde{\boldsymbol{\varepsilon}}) \simeq \mathcal{N}(\tilde{\boldsymbol{\varepsilon}}^r, \tilde{\mathbf{R}}), \quad (1.82)$$

de façon à avoir le couple (\mathbf{x}, \mathbf{y}) gaussien dans le cadre de cette approximation. De nombreuses solutions permettent à présent d'égaliser la moyenne et la matrice de covariance du couple (\mathbf{x}, \mathbf{y}) avec la moyenne et la matrice de covariance du couple $(\mathbf{x}, \tilde{\mathcal{H}}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}})$. Cependant, en imposant $\tilde{\boldsymbol{\varepsilon}}$ de moyenne nulle et décorrélée de \mathbf{x} , on obtient :

$$\tilde{\mathbf{h}} = \mathbb{E}_{p(\mathbf{y})} [\mathbf{y}] = \mathbb{E}_{p(\mathbf{x})} [\mathcal{H}(\mathbf{x})], \quad (1.83a)$$

$$\tilde{\mathbf{H}} = \mathbb{C}_{p(\mathbf{x}, \mathbf{y})} [\mathbf{y}, \mathbf{x}] \mathbb{C}_{p(\mathbf{x})} [\mathbf{x}]^{-1} = \mathbb{C}_{p(\mathbf{x})} [\mathcal{H}(\mathbf{x}), \mathbf{x}] \mathbf{B}^{-1}, \quad (1.83b)$$

$$\tilde{\mathbf{R}} = \mathbb{C}_{p(\mathbf{y})} [\mathbf{y}] - \tilde{\mathbf{H}} \tilde{\mathbf{H}}^T = \mathbf{R} + \mathbb{C}_{p(\mathbf{x})} [\mathcal{H}(\mathbf{x})] - \tilde{\mathbf{H}} \mathbf{B} \tilde{\mathbf{H}}^T. \quad (1.83c)$$

Ainsi défini, $\tilde{\mathcal{H}}$ est donc le meilleur estimateur de \mathbf{y} fonction affine de \mathbf{x} . Puisque nous avons la décomposition :

$$\mathcal{H}(\mathbf{x}) = \tilde{\mathcal{H}}(\mathbf{x}) + \tilde{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}, \quad (1.84)$$

avec $\tilde{\boldsymbol{\varepsilon}} - \boldsymbol{\varepsilon}$ de moyenne nulle et décorrélée de \mathbf{x} , $\tilde{\mathcal{H}}$ est aussi la meilleure approximation affine de \mathcal{H} en erreur quadratique moyenne d'après l'annexe A.3.1. Appliquer les

substitutions $\mathbf{h} \leftarrow \tilde{\mathbf{h}}$, $\mathbf{H} \leftarrow \tilde{\mathbf{H}}$, $\mathbf{R} \leftarrow \tilde{\mathbf{R}}$ aux Eq. (1.65a,1.65b,1.73e,1.73j) :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] \simeq \mathbf{x}^b + \mathbf{B}\tilde{\mathbf{H}}^T (\tilde{\mathbf{H}}\mathbf{B}\tilde{\mathbf{H}}^T + \tilde{\mathbf{R}})^{-1} (\mathbf{y} - \tilde{\mathbf{h}}), \quad (1.85a)$$

$$\simeq \mathbf{x}^b + (\tilde{\mathbf{H}}^T\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{H}} + \mathbf{B}^{-1})^{-1} \tilde{\mathbf{H}}^T\tilde{\mathbf{R}}^{-1} (\mathbf{y} - \tilde{\mathbf{h}}), \quad (1.85b)$$

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}] \simeq \mathbf{B} - \mathbf{B}\tilde{\mathbf{H}}^T (\tilde{\mathbf{H}}\mathbf{B}\tilde{\mathbf{H}}^T + \tilde{\mathbf{R}})^{-1} \tilde{\mathbf{H}}\mathbf{B}, \quad (1.85c)$$

$$\simeq (\tilde{\mathbf{H}}^T\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{H}} + \mathbf{B}^{-1})^{-1}. \quad (1.85d)$$

conduisent alors aux mêmes approximations que Eq. (1.78) d'après Eq. (1.74). Géométriquement, la correction de \mathbf{R} Eq. (1.83c) fait intervenir la courbure de \mathcal{H} . Si l'ensemble a priori \mathbf{E}^b utilisé pour estimer la correction $\mathbb{C}_{p(\mathbf{x})}[\mathcal{H}(\mathbf{x})] - \tilde{\mathbf{H}}\mathbf{B}\tilde{\mathbf{H}}^T$ contient peu de membres ($\leq n + 1$), cette courbure sera indiscernable et $\tilde{\mathbf{R}} = \mathbf{R}$.

Ainsi, l'approximation gaussienne de la densité jointe nous permet d'approcher la moyenne a posteriori par une fonction affine de \mathbf{y} et la covariance a posteriori par une fonction constante de \mathbf{y} . Ce qui est restrictif. Dans la section suivante, nous nous intéressons à l'estimation par maximum a posteriori (MAP). Nous verrons qu'elle constitue un estimateur de \mathbf{x} fonction non linéaire de \mathbf{y} que nous pouvons calculer par des méthodes itératives.

1.1.8 Méthodes variationnelles

Le maximiseur de la densité a posteriori (MAP) ou le minimiseur de la fonction de coût a posteriori est un autre estimateur de \mathbf{x} fonction de \mathbf{y} ne nécessitant pas d'intégration :

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} G(\mathbf{x} = \mathbf{x}|\mathbf{y}). \quad (1.86)$$

Le calcul de \mathbf{x}^* requiert la minimisation de la fonction de coût a posteriori. Contrairement à la méthode section 1.1.7, l'estimateur \mathbf{x}^* de \mathbf{x} n'est donc plus nécessairement une fonction affine de \mathbf{y} . Le maximiseur a posteriori est interprété comme le centre du voisinage infinitésimal le plus probable. La forme « moindre carrés » (Björck, 1996) de la fonction de coût Eq. (1.62) :

$$G(\mathbf{x}|\mathbf{y}) \propto \frac{1}{2} \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathbf{x} - \mathbf{x}^b\|_{\mathbf{B}^{-1}}^2. \quad (1.87)$$

nous indique qu'en un certain sens, le minimum de la fonction de coût minimise l'écart par rapport aux données.

Dans le cas où la fonction de coût a posteriori n'est plus quadratique, l'équation de gradient nul Eq. (1.63a) reste une condition nécessaire d'optimalité. C'est pourquoi

les méthodes d'optimisation numérique cherchent d'abord à résoudre cette équation. Pour vérifier qu'un de ces points singuliers est bien un minimum local strict, il suffit que la matrice hessienne de la fonction de coût en ce point soit symétrique définie-positive. Géométriquement, ces deux conditions signifient qu'au voisinage du minimiseur, la fonction de coût se comporte pratiquement comme une fonction quadratique définie-positive :

$$G(\mathbf{x}|\mathbf{y}) = G(\mathbf{x}^*|\mathbf{y}) + \frac{1}{2} \|\mathbf{x} - \mathbf{x}^*\|_{\nabla_{\mathbf{x}}^2 G(\mathbf{x}^*|\mathbf{y})}^2 + o(\|\mathbf{x} - \mathbf{x}^*\|^2). \quad (1.88)$$

Elles ne garantissent cependant pas le caractère global de \mathbf{x}^* , c'est à dire $\forall \mathbf{x}, G(\mathbf{x}|\mathbf{y}) > G(\mathbf{x}^*|\mathbf{y})$. Ainsi les méthodes présentées ci-dessous peuvent être piégées par un minimum simplement local et, si la symétrie définie-positivité de la matrice hessienne n'est pas vérifiée, par un point singulier (Pires et al., 1996).

Les méthodes d'optimisation numérique du type « line search » (e.g. Nocedal and Wright, 2006) se basent sur les itérations suivantes pour trouver un point singulier :

$$\mathbf{x}^{j+1} = \mathbf{x}^j + \alpha^j \mathbf{p}^j, \quad (1.89)$$

avec α^j la longueur de pas et \mathbf{p}^j la direction de recherche. Les conditions de Wolfe (e.g. Nocedal and Wright, 2006) nous indiquent alors comment choisir ces quantités pour assurer la convergence globale de la méthode, c'est à dire la convergence des itérés $(\mathbf{x}^j)_{j \geq 0}$ vers un point singulier quelque soit le point initial \mathbf{x}^0 . Parmi ces méthodes, on trouve la méthode de la plus grande pente (e.g. Nocedal and Wright, 2006) :

$$\mathbf{p}^j = -\nabla_{\mathbf{x}} G(\mathbf{x}^j|\mathbf{y}). \quad (1.90)$$

Un tel choix garanti $G(\mathbf{x}^{j+1}) < G(\mathbf{x}^j)$ pour des petits pas α^j . Le choix de α^j peut alors se faire par de nombreuses manières comme par exemple, la méthode de « backtracking » (e.g. Nocedal and Wright, 2006). La méthode de la plus grande pente est économique car elle ne requiert que l'évaluation de la dérivée première de la fonction de coût. En revanche, elle converge lentement. A contrario, la méthode de Newton utilise en plus une évaluation de la matrice hessienne pour accélérer cette convergence. L'idée est de former un modèle quadratique de la fonction de coût par développement de Taylor de

second ordre autour de l'itéré courant :

$$G(\mathbf{x}|\mathbf{y}) = J^j(\mathbf{x}|\mathbf{y}) + o\left(\|\mathbf{x} - \mathbf{x}^j\|^2\right), \quad (1.91a)$$

$$J^j(\mathbf{x}|\mathbf{y}) = G(\mathbf{x}^j|\mathbf{y}) + \nabla_{\mathbf{x}}G(\mathbf{x}^j|\mathbf{y})^T (\mathbf{x} - \mathbf{x}^j) + \frac{1}{2} \|\mathbf{x} - \mathbf{x}^j\|_{\nabla_{\mathbf{x}}^2 G(\mathbf{x}^j|\mathbf{y})}^2. \quad (1.91b)$$

Le caractère symétrique défini positif de $\nabla_{\mathbf{x}}^2 G(\mathbf{x}^j)$ est garanti par continuité si \mathbf{x}^j est suffisamment proche de \mathbf{x}^* . L'itéré suivant est alors le minimiseur du modèle, il s'obtient par résolution de l'équation de gradient nul :

$$\mathbf{0} = \nabla_{\mathbf{x}}G(\mathbf{x}^j|\mathbf{y}) + \nabla_{\mathbf{x}}^2 G(\mathbf{x}^j|\mathbf{y}) (\mathbf{x}^{j+1} - \mathbf{x}^j), \Leftrightarrow \quad (1.92a)$$

$$\mathbf{x}^{j+1} = \mathbf{x}^j - \nabla_{\mathbf{x}}^2 G(\mathbf{x}^j|\mathbf{y})^{-1} \nabla_{\mathbf{x}}G(\mathbf{x}^j|\mathbf{y}). \quad (1.92b)$$

En injectant Eq. (1.92b) dans le développement de Taylor de $\nabla_{\mathbf{x}}G$ autour de \mathbf{x}^j , on obtient la convergence quadratique des itérés vers un point singulier :

$$\nabla_{\mathbf{x}}G(\mathbf{x}^*|\mathbf{y}) = \mathbf{0} = \nabla_{\mathbf{x}}G(\mathbf{x}^j|\mathbf{y}) + \nabla^2 G(\mathbf{x}^j|\mathbf{y}) (\mathbf{x}^j - \mathbf{x}^*) + o\left(\|\mathbf{x}^j - \mathbf{x}^*\|^2\right), \Leftrightarrow \quad (1.93a)$$

$$\mathbf{x}^{j+1} - \mathbf{x}^* = o\left(\|\mathbf{x}^j - \mathbf{x}^*\|^2\right). \quad (1.93b)$$

L'existence, pour chaque minimum local stricte, d'un voisinage où la fonction de coût est strictement convexe assure donc la convergence locale de la méthode. Pour garder une vitesse de convergence raisonnable tout en évitant l'évaluation coûteuse de la matrice hessienne, la méthode de Gauss-Newton exploite la forme « moindres carrés » de la fonction de coût Eq. (1.87). L'idée est de remplacer $\mathcal{H}(\mathbf{x})$ par son développement de Taylor de premier ordre autour de l'itéré courant :

$$\mathcal{H}(\mathbf{x}) \simeq \mathbf{h}^j + \mathbf{H}^j (\mathbf{x} - \mathbf{x}^j), \quad (1.94a)$$

$$\begin{bmatrix} \mathbf{h}^j & \mathbf{H}^j \end{bmatrix} = \begin{bmatrix} \mathcal{H}(\mathbf{x}^j) & \frac{d\mathcal{H}}{d\mathbf{x}}(\mathbf{x}^j) \end{bmatrix}. \quad (1.94b)$$

afin de former un modèle quadratique $J^j(\mathbf{x})$ de $G(\mathbf{x})$:

$$G(\mathbf{x}|\mathbf{y}) = J^j(\mathbf{x}|\mathbf{y}) + o\left(\|\mathbf{x} - \mathbf{x}^j\|\right),$$

$$J^j(\mathbf{x}|\mathbf{y}) = \frac{1}{2} \|\mathbf{x} - \mathbf{x}^b\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \|\mathbf{y} - \mathbf{h}^j - \mathbf{H}^j (\mathbf{x} - \mathbf{x}^j)\|_{\mathbf{R}^{-1}}^2, \quad (1.95a)$$

L'itéré suivant est aussi le minimiseur du modèle. Il s'obtient par résolution de l'équation

de gradient nul pour J^j :

$$\nabla_{\mathbf{x}} J^j(\mathbf{x}^{j+1}|\mathbf{y}) = \mathbf{0}, \Leftrightarrow \quad (1.96a)$$

$$\nabla_{\mathbf{x}} J^j(\mathbf{x}^j|\mathbf{y}) + \nabla_{\mathbf{x}}^2 J^j(\mathbf{x}^{j+1} - \mathbf{x}^j) = \mathbf{0}, \Leftrightarrow \quad (1.96b)$$

$$\mathbf{x}^{j+1} = \mathbf{x}^j - \mathbf{C}^{j+1} \left(\mathbf{B}^{-1}(\mathbf{x}^j - \mathbf{x}^b) - (\mathbf{H}^j)^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}^j) \right) \quad (1.96c)$$

$$\mathbf{C}^{j+1} = \left(\mathbf{B}^{-1} + (\mathbf{H}^j)^\top \mathbf{R}^{-1} (\mathbf{H}^j) \right)^{-1}, \quad (1.96d)$$

A cause de la différence des hessiennes $\nabla_{\mathbf{x}}^2 J^j \neq \nabla_{\mathbf{x}}^2 G(\mathbf{x}^j|\mathbf{y})$, un ordre est perdu dans l'approximation de G aussi bien que la convergence locale. Cependant, la symétrie définie-positivité de $\nabla_{\mathbf{x}}^2 J^j$ est garantie, ce qui rend l'algorithme plus stable et efficace en pratique.

L'estimation par maximum a posteriori souffre en revanche d'un défaut majeur ; il ne fournit pas naturellement d'estimateur de second ordre permettant d'évaluer l'erreur d'estimation. On peut néanmoins s'intéresser aux situations dans lesquelles le modèle Eq. (1.95a) a une pertinence statistique :

$$p(\mathbf{x}|\mathbf{y}) \simeq p^j(\mathbf{x}|\mathbf{y}), \quad (1.97a)$$

$$p^j(\mathbf{x}|\mathbf{y}) \propto e^{-J^j(\mathbf{x}|\mathbf{y})}. \quad (1.97b)$$

L'approximation de Laplace (Di Persio, 2018) permet, dans un cadre asymptotique, de faire ce lien. L'idée est de remarquer que l'exponentielle dans la densité Eq. (1.97b) concentre fortement la masse de probabilité dans une zone où la fonction de coût $G(\mathbf{x}|\mathbf{y})$ est petite. Cette concentration de la masse de probabilité autour de \mathbf{x}^* peut être suffisante pour que l'approximation affine locale Eq. (1.94a) soit précise dans cette zone de forte probabilité. Dans ce cas, $p(\mathbf{x}|\mathbf{y})$ se comporte comme la gaussienne $p^j(\mathbf{x}|\mathbf{y})$. En effet, si ces densités sont proches dans une zone de forte probabilité, l'espérance d'une fonction réelle $\varphi(\mathbf{x})$ quelconque selon $p(\mathbf{x}|\mathbf{y})$ est relativement proche de son intégration selon $p^j(\mathbf{x}|\mathbf{y})$. Pour illustrer cela, on peut appliquer l'inégalité de Cauchy-Schwartz à l'erreur d'intégration :

$$\left| \mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\varphi(\mathbf{x})] - \mathbb{E}_{p^j(\mathbf{x}|\mathbf{y})}[\varphi(\mathbf{x})] \right| = \left| \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[\varphi(\mathbf{x}) \left(1 - \frac{p^j(\mathbf{x}|\mathbf{y})}{p(\mathbf{x}|\mathbf{y})} \right) \right] \right|, \quad (1.98a)$$

$$\leq \mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\varphi^2(\mathbf{x})]^{1/2} \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[\left(1 - \frac{p^j(\mathbf{x}|\mathbf{y})}{p(\mathbf{x}|\mathbf{y})} \right)^2 \right]^{1/2}. \quad (1.98b)$$

L'erreur relative est alors bornée par une quantité ne dépendant pas de φ et qui est faible si l'approximation est précise dans une zone de forte probabilité :

$$\frac{|\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\varphi(\mathbf{x})] - \mathbb{E}_{p^j(\mathbf{x}|\mathbf{y})} [\varphi(\mathbf{x})]|}{\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\varphi^2(\mathbf{x})]^{\frac{1}{2}}} \leq \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[\left(1 - \frac{p^j(\mathbf{x}|\mathbf{y})}{p(\mathbf{x}|\mathbf{y})} \right)^2 \right]^{\frac{1}{2}}, \quad (1.99)$$

le membre de droite est aussi la divergence du chi carré de Pearson (qui est proche de l'entropie relative, [Cover and Thomas, 2005](#)). Dans ce cas, on fera les approximations :

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}] \simeq \mathbf{x}^{j+1}, \quad (1.100a)$$

$$\mathbb{C}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}] \simeq \mathbf{C}^{j+1}. \quad (1.100b)$$

Lorsque la densité de probabilité n'est pas suffisamment concentrée, on peut reprocher au développement limité de \mathcal{H} Eq. (1.94a) d'être trop local, sa validité n'étant garantie que pour un voisinage infinitésimal de \mathbf{x}^j . Une alternative s'inspirant de Eq. (1.83) est de choisir une approximation affine \mathcal{H}^j de \mathcal{H} minimisant les erreurs quadratiques moyennes selon $p^{j-1}(\mathbf{x}|\mathbf{y})$:

$$\mathbb{E}_{p^{j-1}(\mathbf{x}|\mathbf{y})} \left[\left\| \mathcal{H}(\mathbf{x}) - \mathcal{H}^j(\mathbf{x}) \right\|^2 \right]. \quad (1.101)$$

Puisque la substitution $\mathcal{H} \leftarrow \mathcal{H}^j$ dans l'expression de $G(\mathbf{x}|\mathbf{y})$ Eq. (1.87) engendre une nouvelle approximation gaussienne $p^j(\mathbf{x}|\mathbf{y})$ de $p(\mathbf{x}|\mathbf{y})$, la méthode peut être itérée. Avec $p^0(\mathbf{x}|\mathbf{y}) = p(\mathbf{x})$ (et la correction $\tilde{\mathbf{R}}$ Eq. (1.83c) de la matrice de covariance d'erreur \mathbf{R}), la première itération de la méthode donne les mêmes estimations des moments a posteriori que celles de la section 1.1.7. On quitte alors la méthode de Gauss-Newton pour se rapprocher des méthodes de quadrature itératives (e.g. [Bernardo and Smith, 2008](#)).

1.1.9 Conclusion de l'assimilation de données sans dimension temporelle

Nous avons vu dans cette section des techniques permettant d'analyser l'état \mathbf{x} à l'aide de l'assimilation de l'observation \mathbf{y} . Dans le cadre bayésien, la solution de cette analyse est la densité a posteriori $p(\mathbf{x}|\mathbf{y})$. La règle de Bayes permet d'exprimer cette densité a posteriori en fonction de la vraisemblance et de la densité a priori.

Cependant, cette expression est de faible intérêt en pratique, car l'intégration nécessaire au calcul de ses moments n'est pas calculable. L'alternative proposée par l'échan-

tillonnage d'importance est d'approcher ces intégrales par des formules de quadrature. Ce qui revient à approcher les densités par des sommes de Dirac. Bien qu'exacte dans la limite du nombre d'échantillons, cette méthode requiert un grand nombre d'évaluations de l'opérateur d'observation \mathcal{H} . Puisque cet opérateur porte l'essentiel de la complexité numérique, la méthode par échantillonnage d'importance n'est pas envisageable en grande dimension.

En revanche, l'opérateur d'observation \mathcal{H} est affine si et seulement si la densité jointe $p(\mathbf{x}, \mathbf{y})$ est gaussienne. La moyenne et la matrice de covariance a posteriori caractérisent alors la densité a posteriori et peuvent être calculées algébriquement. La moyenne a posteriori est, dans ce cas, le meilleur estimateur de \mathbf{x} fonction de \mathbf{y} en erreur quadratique moyenne (MMSE), le meilleur estimateur de \mathbf{x} fonction affine de \mathbf{y} en erreur quadratique moyenne (BLUE), ainsi que la maximum a posteriori (MAP). Dans le cas général, on peut s'inspirer du cas linéaire pour faire une approximation gaussienne de $p(\mathbf{x}|\mathbf{y})$ ne retenant que ses deux premiers moments. La moyenne et la matrice de covariance a posteriori ne caractérisent plus la densité a posteriori, elles engendrent néanmoins sa meilleure approximation gaussienne au sens de l'entropie relative. En particulier, la moyenne a posteriori est le meilleur estimateur de \mathbf{x} fonction de \mathbf{y} au sens de l'erreur quadratique moyenne (MMSE). Le calcul de ces moments a posteriori demande cependant une intégration par rapport à $p(\mathbf{x}|\mathbf{y})$. Cette intégration est impossible parce que l'intégrande $p(\mathbf{x}|\mathbf{y})$ est justement la densité que l'on cherche à calculer. En revanche, si l'on sait intégrer $(\mathbf{x}, \mathcal{H}(\mathbf{x}))$ par rapport à la densité a priori $p(\mathbf{x})$, la moyenne et la matrice de covariance de (\mathbf{x}, \mathbf{y}) sont accessibles et fournissent la meilleure approximation gaussienne de la densité jointe $p(\mathbf{x}, \mathbf{y})$ au sens de l'entropie relative. Cette approximation gaussienne de $p(\mathbf{x}, \mathbf{y})$ correspond à une approximation affine de \mathcal{H} minimisant l'erreur quadratique moyenne avec une approximation gaussienne de son erreur. On obtient alors une approximation de moments a posteriori dont l'approximation de la moyenne a posteriori est le meilleur estimateur de \mathbf{x} fonction affine de \mathbf{y} en erreur quadratique moyenne.

Finalement, le MAP de $p(\mathbf{x}|\mathbf{y})$ fournit un estimateur fonction non linéaire de \mathbf{y} calculable grâce aux méthodes itératives d'optimisation. Cette technique ne fournit malheureusement pas naturellement d'estimateur de second ordre. Une alternative consiste alors à donner une portée statistique à l'approximation affine de \mathcal{H} au sein de la méthode de Gauss-Newton. Cette approximation se doit d'être précise dans les zones de forte probabilité pour que le calcul des moments soit relativement bon. A leur tour, ces moments peuvent servir à mieux localiser la zone de forte probabilité et ainsi améliorer notre approximation affine de \mathcal{H} .

Dans la section suivante, nous nous intéresserons à l'assimilation de données avec

une dimension temporelle. Les observations sont alors disponibles à mesure que le temps passe.

1.2 L'assimilation de données avec dimension temporelle

Le système est à présent enrichi d'une dimension temporelle :

$$\mathbf{y}_l = \mathcal{H}_l(\mathbf{x}_l) + \boldsymbol{\varepsilon}_l^r, \quad (1.102a)$$

$$\mathbf{x}_l = \mathcal{M}_l(\mathbf{x}_{l-1}) + \boldsymbol{\varepsilon}_l^q, \quad (1.102b)$$

avec $l \in \mathbb{Z}$. À chaque instant t_l correspond une observation \mathbf{y}_l et un état \mathbf{x}_l . La dynamique du système est décrite par un ensemble d'équations différentielles dont la résolvente \mathcal{M}_l fait évoluer l'état \mathbf{x}_{l-1} vers l'état \mathbf{x}_l . Puisque la résolvente \mathcal{M}_l est elle aussi une représentation simplifiée de la réalité, elle est nécessairement inexacte. Ce que représente l'erreur modèle $\boldsymbol{\varepsilon}_l^q$. On supposera les erreurs d'observation et de modèle $(\boldsymbol{\varepsilon}_l^r, \boldsymbol{\varepsilon}_l^q)_{l \in \mathbb{Z}}$ mutuellement indépendantes avec pour densité :

$$p(\boldsymbol{\varepsilon}_l^r) = \mathcal{N}(\boldsymbol{\varepsilon}_l^r | \mathbf{0}, \mathbf{R}_l), \quad (1.103a)$$

$$p(\boldsymbol{\varepsilon}_l^q) = \mathcal{N}(\boldsymbol{\varepsilon}_l^q | \mathbf{0}, \mathbf{Q}_l). \quad (1.103b)$$

Les matrices de covariances d'erreur d'observation \mathbf{R}_l et d'erreur modèle \mathbf{Q}_l sont supposées symétriques, définies positives.

1.2.1 Solution bayésienne

La solution du problème d'assimilation est toujours la densité a posteriori. Cependant celle-ci dépend à présent d'une fenêtre d'assimilation (e.g. [Carrassi et al., 2017a](#)) :

$$p(\mathbf{x}_{0:L} | \mathbf{y}_{:L}), \quad (1.104)$$

où t_0 est défini comme l'origine de la fenêtre d'assimilation courante, le paramètre L est ainsi la longueur de fenêtre. Lorsqu'elle s'applique à des vecteurs ou à des ensembles, la tranche $0 : L$ correspond à la concaténation verticale des vecteurs aléatoires selon le

temps :

$$\mathbf{x}_{0:L} = \begin{bmatrix} \mathbf{x}_0 \\ \vdots \\ \mathbf{x}_L \end{bmatrix}. \quad (1.105)$$

La tranche : L représente tous les indices temporels depuis le début de l'assimilation. Lorsque $L = 0$ il s'agit de la solution du problème de filtrage, sinon il s'agit de la solution du problème de lissage (Tukey and Wiener, 1952). Le problème de lissage est donc plus général que le problème de filtrage puisque la solution du second s'obtient par marginalisation de la solution du premier.

Contrairement à l'assimilation sans dimension temporelle, on ne calculera pas la solution en une fois. Avec la dimension temporelle, l'assimilation peut être faite de manière séquentielle. Dans ce cadre, la solution du cycle courant $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ utilise la solution du cycle précédent $p(\mathbf{x}_{-S:L-S}|\mathbf{y}_{:L-S})$ avec S le nombre de pas de temps dont la fenêtre est translétée d'un cycle à un autre. Si les nouvelles observations à assimiler au cycle courant sont $\mathbf{y}_{K:L}$, il est nécessaire d'avoir $L - S = K - 1$ pour que chaque observation ne soit assimilée qu'une seule fois. Une représentation de la fenêtre d'assimilation est donnée Fig. (1.2)

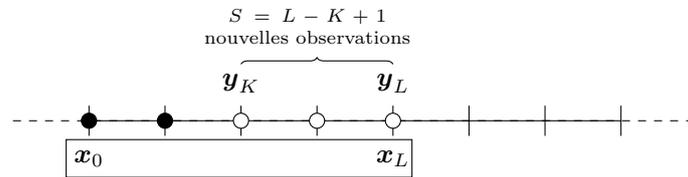


FIGURE 1.2 – Une fenêtre d'assimilation paramétrée par $K = 2$, $L = 4$ et $S = 3$. Les points noirs représentent les observations assimilées au cycle précédent ; les points blancs sont les nouvelles observations assimilées au cycle courant ; les observations sans point ne sont pas encore assimilées ; le cadre regroupe les variables d'état de la fenêtre.

La règle de Bayes s'écrit alors :

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) = \frac{p(\mathbf{y}_{K:L}|\mathbf{x}_{0:L}, \mathbf{y}_{:K-1})}{p(\mathbf{y}_{K:L}|\mathbf{y}_{:K-1})} p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}), \quad (1.106a)$$

$$\propto p(\mathbf{y}_{K:L}|\mathbf{x}_{K:L}) p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}). \quad (1.106b)$$

La simplification $p(\mathbf{y}_{K:L}|\mathbf{x}_{0:L}, \mathbf{y}_{:K-1}) \rightarrow p(\mathbf{y}_{K:L}|\mathbf{x}_{K:L})$ provient de l'indépendance des

erreurs. La densité a priori est ici $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S})$ puisqu'elle ne contient pas les nouvelles observations $\mathbf{y}_{K:L}$. Son expression (ou une approximation) est supposée disponible par récurrence. Le calcul de la densité a posteriori $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ (ou une approximation) à partir de la densité a priori $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S})$ constitue l'étape d'analyse. Cette étape est illustrée Fig. (1.3).

La densité a priori du cycle suivant $p(\mathbf{x}_{S:L+S}|\mathbf{y}_{:L})$ s'obtient ensuite en deux phases. La première étend la fenêtre d'assimilation dans le futur pour prendre en compte les nouveaux états $\mathbf{x}_{L+1:L+S}$:

$$p(\mathbf{x}_{0:L+S}|\mathbf{y}_{:L}) = p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_{0:L}, \mathbf{y}_{:L}) p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}), \quad (1.107a)$$

$$= p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_L) p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}). \quad (1.107b)$$

La simplification $p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_{0:L}, \mathbf{y}_{:L}) \rightarrow p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_L)$ provient de l'indépendance des erreurs. La seconde phase réduit la fenêtre d'assimilation dans le passé par marginalisation pour éliminer les anciens états $\mathbf{x}_{0:S-1}$:

$$p(\mathbf{x}_{S:L+S}|\mathbf{y}_{:L}) = \int p(\mathbf{x}_{0:L+S}|\mathbf{y}_{:L}) d\mathbf{x}_{0:S-1}. \quad (1.108)$$

La succession de ces deux phases constitue l'étape de propagation. Elle est illustrée Fig. (1.4).

En résumé, un cycle d'assimilation est la succession de ces étapes d'analyse et de propagation :

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto p(\mathbf{y}_{K:L}|\mathbf{x}_{K:L}) p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}), \quad (\text{Analyse}) \quad (1.109a)$$

$$p(\mathbf{x}_{S:L+S}|\mathbf{y}_{:L}) = \int p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_L) p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) d\mathbf{x}_{0:S-1}, \quad (\text{Propagation}) \quad (1.109b)$$

Dans l'étape d'analyse on ajoute de nouvelles observations, dans l'étape de propagation on translate la fenêtre d'assimilation.

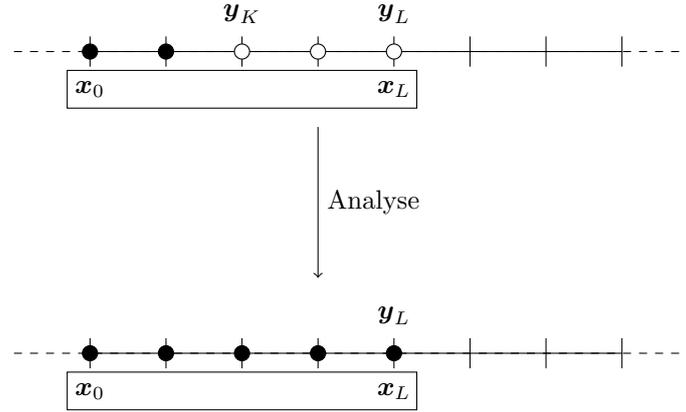


FIGURE 1.3 – L'étape d'analyse, les nouvelles observations $\mathbf{y}_{L:K}$ sont assimilées.

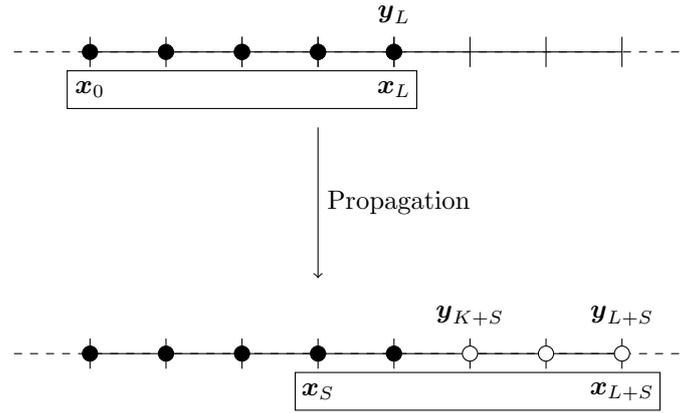


FIGURE 1.4 – L'étape de propagation, la fenêtre d'assimilation est translatée de S pas de temps.

Des changements de variables et l'indépendance des erreurs d'observation permettent de préciser l'expression Eq. (1.109a) :

$$\begin{aligned}
 p(\mathbf{x}_{0:L} = \mathbf{x}_{0:L} | \mathbf{y}_{:L} = \mathbf{y}_{:L}) &\propto \prod_{l=K}^L p(\boldsymbol{\varepsilon}_l = \mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)) \\
 &\times p(\mathbf{x}_{0:L} = \mathbf{x}_{0:L} | \mathbf{y}_{:L-S} = \mathbf{y}_{:L-S}), \quad (1.110a)
 \end{aligned}$$

En remplaçant les densités $p(\boldsymbol{\varepsilon}_l)$, $p(\boldsymbol{\eta}_l)$ par les gaussiennes Eq. (1.103a,1.103b) on ob-

tient, par passage au logarithme, les fonctions de coût exactes :

$$G(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto \sum_{l=K}^L \frac{1}{2} \|\mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)\|_{\mathbf{R}_l^{-1}}^2 + G(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}), \quad (1.111)$$

où la notation \propto signifie « égal à une constante additive près » lorsqu'elle s'applique à des fonctions de coût.

Dans le cas fortement contraint, c'est à dire lorsque les erreurs modèle $(\boldsymbol{\varepsilon}_l^q)_l$ sont négligées, la trajectoire $\mathbf{x}_{0:L}$ est déterminée par sa condition initiale \mathbf{x}_0 avec :

$$\mathbf{x}_l = \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0), \quad (1.112)$$

pour $l = 0, \dots, L$. Dans ce cas, la densité $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ est ramenée à la densité $p(\mathbf{x}_0|\mathbf{y}_{:L})$:

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto p(\mathbf{x}_0|\mathbf{y}_{:L}) \prod_{l=1}^L \delta(\mathbf{x}_l - \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0)), \quad (1.113)$$

qui a pour fonction de coût associée :

$$G(\mathbf{x}_0|\mathbf{y}_{:L}) \propto \sum_{l=K}^L \frac{1}{2} \|\mathbf{y}_l - \mathcal{H}_l \circ \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0)\|_{\mathbf{R}_l^{-1}}^2 + G(\mathbf{x}_0|\mathbf{y}_{:L-S}), \quad (1.114)$$

Ces expressions analytiques permettent en théorie d'exprimer séquentiellement la densité a posteriori. Cependant, elles n'ont que peu d'intérêt en pratique car il est difficile de les manipuler. Lorsque la dynamique du modèle est chaotique et instable, les observations passées ont peu d'importance pour estimer les états de la fenêtre courante d'assimilation (Bocquet and Carrassi, 2017). C'est pourquoi on peut se permettre de simplifier la fonction de coût a priori $G(\mathbf{x}_0|\mathbf{y}_{:L-S})$ à l'aide d'une approximation. C'est cette approximation qui permet aux méthodes séquentielles d'assimilation de données d'être efficaces en pratique sur de grandes périodes temporelles.

1.2.2 Le filtre particulaire

Le filtre particulaire (Doucet and Johansen, 2011; Gordon et al., 1993) est une méthode d'assimilation de données approchant la solution bayésienne du problème de filtrage dans un cadre séquentiel sans reposer sur la linéarité des opérateurs. Dans la fenêtre d'assimilation $L = 0$, on suivra ici la succession d'étapes bayésiennes :

$$p(\mathbf{x}_0|\mathbf{y}_{:-1}) \xrightarrow{\text{A}} p(\mathbf{x}_0|\mathbf{y}_{:0}) \xrightarrow{\text{P}} p(\mathbf{x}_1|\mathbf{y}_{:0}). \quad (1.115)$$

L'approche particulière permet ainsi d'estimer les moments de ces densités à l'aide d'ensembles.

Analyse : elle consiste à engendrer un ensemble $\mathbf{E}_0^a = (\mathbf{x}_0^{1,a}, \dots, \mathbf{x}_0^{M,a})$ et des poids $\mathbf{w}_0^a = (w_0^{1,a}, \dots, w_0^{M,a})^\top$ échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:0})$ à partir d'un ensemble $\mathbf{E}_0^b = (\mathbf{x}_0^{1,b}, \dots, \mathbf{x}_0^{M,b})$ et des poids $\mathbf{w}_0^b = (w_0^{1,b}, \dots, w_0^{M,b})^\top$ échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{-1})$. On procède pour cela comme dans la section 1.1.3. L'analogie de la solution Eq. (1.29) est alors :

$$w_0^{i,a} \propto p(\mathbf{y}_0|\mathbf{x}_0^{i,b}) w_0^{i,b}, \quad (1.116a)$$

$$\mathbf{E}_0^a = \mathbf{E}_0^b. \quad (1.116b)$$

Pour calculer la constante normalisatrice, on impose aussi que la formule de quadrature Eq. (1.25a) soit exacte pour les fonctions φ constantes.

Propagation : elle consiste à engendrer un ensemble \mathbf{E}_1^b et des poids \mathbf{w}_1^b échantillonnant $p(\mathbf{x}_1|\mathbf{y}_{:0})$ à partir de l'ensemble \mathbf{E}_0^a et des poids \mathbf{w}_0^a échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:0})$. D'après Eq. (1.102b) et l'indépendance des erreurs, on a pour toute fonction φ :

$$\mathbb{E}_{p(\mathbf{x}_1|\mathbf{y}_{:0})} [\varphi(\mathbf{x}_1)] = \mathbb{E}_{p(\mathbf{x}_0, \boldsymbol{\varepsilon}_1^q|\mathbf{y}_{:0})} [\varphi(\mathcal{M}(\mathbf{x}_0) + \boldsymbol{\varepsilon}_1^q)], \quad (1.117a)$$

$$= \mathbb{E}_{p(\boldsymbol{\varepsilon}_1^q)} \left[\mathbb{E}_{p(\mathbf{x}_0|\mathbf{y}_{:0})} [\varphi(\mathcal{M}(\mathbf{x}_0) + \boldsymbol{\varepsilon}_1^q)] \right], \quad (1.117b)$$

$$\simeq \sum_{j=1}^{M'} \frac{1}{M'} \sum_{i=1}^M w_0^{i,a} \varphi(\mathcal{M}(\mathbf{x}_0^{i,a}) + \boldsymbol{\varepsilon}_1^{j,q}), \quad (1.117c)$$

où $(\boldsymbol{\varepsilon}_1^{1,q}, \dots, \boldsymbol{\varepsilon}_1^{M',q})$ est un ensemble associé au poids constant $\frac{1}{M'}$ échantillonnant $\mathcal{N}(\boldsymbol{\varepsilon}_1^q|\mathbf{0}_n, \mathbf{Q}_1)$. Une solution est donc un ensemble \mathbf{E}_1^b et des poids \mathbf{w}_1^b de taille $M \times M'$ tels que :

$$\mathbf{x}_1^{i,j,b} = \mathcal{M}(\mathbf{x}_0^{i,a}) + \boldsymbol{\varepsilon}_1^{j,q}, \quad (1.118a)$$

$$w^{i,j,b} = \frac{w^{i,a}}{M'}, \quad (1.118b)$$

pour $(i, j) \in \{1, \dots, M\} \times \{1, \dots, M'\}$. Cependant, la taille de l'ensemble a été multipliée par M' . Pour éviter cela, une solution simple consiste à ne conserver que les éléments diagonaux $i = j$ avec $M = M'$ et à normaliser les poids. On obtient ainsi :

$$\mathbf{x}_1^{i,b} = \mathcal{M}(\mathbf{x}_0^{i,a}) + \boldsymbol{\varepsilon}_1^{i,q}, \quad (1.119a)$$

$$w^{i,b} = w^{i,a}, \quad (1.119b)$$

pour $i = 1, \dots, M$.

D'un point de vue théorique, le filtre particulaire résout exactement le problème de filtrage lorsque le nombre de membres M des ensembles tend vers l'infini. En revanche, le filtre particulaire se révèle instable en pratique. En effet, un poids a tendance à concentrer toute la masse alors que les autres sont proches de zéro (Kong and Liu, 1994). Pour éviter cette dégénérescence, il faut alors augmenter le nombre de points de manière exponentielle avec la dimension de l'espace des états (Snyder et al., 2008). Le filtre particulaire est donc adapté aux problèmes d'assimilation de données hautement non linéaire mais de faible dimension. Cette contrainte est souvent appelée « la malédiction de la dimension ».

1.2.3 Solution dans le cadre linéaire : le filtre de Kalman

Le filtre de Kalman est la solution séquentielle du problème de filtrage dans le cas d'opérateurs affines et d'erreurs gaussiennes. Dans la fenêtre d'assimilation $L = 0$, on suivra ici exactement la succession d'étapes bayésiennes :

$$p(\mathbf{x}_0 | \mathbf{y}_{:-1}) \xrightarrow{A} p(\mathbf{x}_0 | \mathbf{y}_{:0}) \xrightarrow{P} p(\mathbf{x}_1 | \mathbf{y}_{:0}). \quad (1.120)$$

grâce à la linéarité des opérateurs \mathcal{H}_0 et \mathcal{M}_1 :

$$\mathcal{H}_0(\mathbf{x}_0) = \mathcal{H}_0(\mathbf{x}_0^b) + \mathbf{H}_0(\mathbf{x}_0 - \mathbf{x}_0^b), \quad (1.121a)$$

$$\mathcal{M}_1(\mathbf{x}_0) = \mathcal{M}_1(\mathbf{x}_0^b) + \mathbf{M}_1(\mathbf{x}_0 - \mathbf{x}_0^b). \quad (1.121b)$$

En effet, avec ces hypothèses, les étapes d'analyse et de propagation conservent le caractère gaussien des densités. Puisque les densités gaussiennes sont caractérisées par leurs moyennes et leurs matrices de covariance, chacune de ces étapes se traduit donc en termes d'opérations algébriques sur ces moments. La moyenne a posteriori calculée par le filtre de Kalman est donc le meilleur estimateur de \mathbf{x} sachant $\mathbf{y}_{:0}$ en erreur quadratique moyenne. Pour d'autres constructions du filtre de Kalman, on pourra consulter (Bain and Crisan, 2009; Bengtsson et al., 1981; Cohn, 1997; Jazwinski, 1970; Kalman, 1960; Talagrand, 1997; Wikle and Berliner, 2007).

Analyse : elle consiste à calculer la moyenne \mathbf{x}_0^a et la matrice de covariance \mathbf{C}_0^a de $p(\mathbf{x}_0 | \mathbf{y}_{:0})$ à partir de la moyenne \mathbf{x}_0^b et la matrice de covariance \mathbf{B}_0 de $p(\mathbf{x}_0 | \mathbf{y}_{:-1})$. Par hypothèse de récurrence, on supposera donc que la fonction de coût a priori est de la

forme :

$$G(\mathbf{x}_0|\mathbf{y}_{:-1}) \propto \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0^{-1}}^2, \quad (1.122)$$

avec \mathbf{B}_0 symétrique, définie positive. Dans ce cas, $G(\mathbf{x}_0|\mathbf{y}_{:0})$ Eq. (1.111) est une fonction quadratique :

$$G(\mathbf{x}_0|\mathbf{y}_{:0}) \propto \frac{1}{2} \|\mathbf{y}_0 - \mathcal{H}_0(\mathbf{x}_0)\|_{\mathbf{R}_0^{-1}}^2 + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0^{-1}}^2. \quad (1.123)$$

donc $p(\mathbf{x}_0|\mathbf{y}_{:0})$ est une densité gaussienne. Elle est alors caractérisée par sa moyenne \mathbf{x}_0^a et sa matrice de covariance \mathbf{C}_0^a . L'analyse consiste alors à calculer ces moments à l'aide de l'algèbre de la section 1.1.6 :

$$\mathbf{x}_0^a = \mathbf{x}_0^b + \mathbf{C}_0^a \mathbf{H}_0^T \mathbf{R}_0^{-1} (\mathbf{y}_0 - \mathcal{H}(\mathbf{x}_0^b)), \quad (1.124a)$$

$$\mathbf{C}_0^a = (\mathbf{B}_0^{-1} + \mathbf{H}_0^T \mathbf{R}_0^{-1} \mathbf{H}_0)^{-1}. \quad (1.124b)$$

Propagation : elle consiste à calculer la moyenne \mathbf{x}_1^b et la matrice de covariance \mathbf{B}_1 de $p(\mathbf{x}_1|\mathbf{y}_{:0})$ à partir de la moyenne \mathbf{x}_0^a et de la matrice de covariance \mathbf{C}_0^a de $p(\mathbf{x}_0|\mathbf{y}_{:0})$. D'après Eq. (1.102b), du caractère affine du modèle et du caractère gaussien de $p(\mathbf{x}_0, \boldsymbol{\varepsilon}_1^q|\mathbf{y}_{:0}) = p(\mathbf{x}_0|\mathbf{y}_{:0})p(\boldsymbol{\varepsilon}_1^q)$, la densité a priori du prochain cycle $p(\mathbf{x}_1|\mathbf{y}_{:0})$ est gaussienne. Ces moments sont :

$$\mathbf{x}_1^b = \mathcal{M}_1(\mathbf{x}_0^a), \quad (1.125a)$$

$$\mathbf{B}_1 = \mathbf{M}_1 \mathbf{C}_0^a \mathbf{M}_1^T + \mathbf{Q}_1. \quad (1.125b)$$

Le filtre de Kalman n'est malheureusement pas applicable aux systèmes de grande taille pour plusieurs raisons. La première est qu'il ne gère pas des opérateurs non linéaires, ce qui est rédhibitoire pour la plupart des applications physiques réalistes. La seconde est qu'il n'est pas efficace en grande dimension. En effet, le stockage d'une matrice de covariance requiert $n(n+1)/2$ scalaires. En géophysique, ceci nécessite une trop grande quantité de mémoire. Aussi, l'inversion matricielle Eq.(1.124b) est coûteuse en nombre d'opérations. Si la dimension d de l'espace des observations est petite devant n , ce problème peut être contourné à l'aide des formules de Sherman-Morrison-Woodbury. En revanche, les deux produits matriciels Eq. (1.125b) restent prohibitifs.

L'enjeu de l'assimilation de données séquentielle est donc de construire un algorithme précis même lorsque les opérateurs sont non linéaires et efficace en termes d'opérations numériques et d'utilisation de la mémoire.

1.2.4 Solution dans le cadre linéaire : le lisseur de Kalman

Le lisseur de Kalman généralise le filtre de Kalman puisqu'il s'agit de calculer séquentiellement la solution du problème de lissage toujours dans le cas d'opérateurs affines. Dans la fenêtre d'assimilation $0 \leq K \leq L$, on suivra donc ici exactement la succession d'étapes bayésienne :

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}) \xrightarrow{A} p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \xrightarrow{P} p(\mathbf{x}_{S:L+S}|\mathbf{y}_{:L}). \quad (1.126)$$

Par rapport au filtre de Kalman, on passe de l'estimation d'un état \mathbf{x}_0 à l'estimation d'une trajectoire $\mathbf{x}_{0:L}$. Pour d'autres constructions du lisseur de Kalman on pourra consulter [Jazwinski \(1970\)](#), [Cosme et al. \(2012\)](#).

Analyse : elle consiste à calculer la moyenne $\mathbf{x}_{0:L}^a$ et la matrice de covariance $\mathbf{C}_{0:L}^a$ de $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ à partir de la moyenne $\mathbf{x}_{0:L}^b$ et la matrice de covariance $\mathbf{B}_{0:L}$ de $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S})$. Par hypothèse de récurrence, on supposera que cette dernière densité a priori est la gaussienne :

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}) \propto e^{-G(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S})}, \quad (1.127a)$$

$$G(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}) \propto \frac{1}{2} \|\mathbf{x}_{0:L} - \mathbf{x}_{0:L}^b\|_{\mathbf{B}_{0:L}^{-1}}^2, \quad (1.127b)$$

avec $\mathbf{B}_{0:L}$ symétrique, définie positive. Dans ce cas, $G(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ Eq. (1.111) est une fonction quadratique, définie positive :

$$G(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto \frac{1}{2} \sum_{l=K}^L \|\mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)\|_{\mathbf{R}_l^{-1}}^2 + \frac{1}{2} \|\mathbf{x}_{0:L} - \mathbf{x}_{0:L}^b\|_{\mathbf{B}_{0:L}^{-1}}^2. \quad (1.128)$$

donc $p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})$ est une densité gaussienne. Elle est alors caractérisée par sa moyenne $\mathbf{x}_{0:L}^a$ et sa matrice de covariance $\mathbf{C}_{0:L}^a$. L'analyse consiste alors à calculer ces moments à l'aide de l'algèbre section 1.1.6. Pour simplifier les calculs, on peut mettre la fonction de coût Eq. (1.128) sous une forme 4D :

$$G(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto \frac{1}{2} \|\mathbf{y}_{K:L} - \mathcal{H}_{K:L}(\mathbf{x}_{0:L})\|_{\mathbf{R}_{K:L}^{-1}}^2 + \frac{1}{2} \|\mathbf{x}_{0:L} - \mathbf{x}_{0:L}^b\|_{\mathbf{B}_{0:L}^{-1}}^2, \quad (1.129)$$

avec

$$\mathcal{H}_{K:L}(\mathbf{x}_{0:L}) = \begin{bmatrix} \mathcal{H}_K(\mathbf{x}_K) \\ \vdots \\ \mathcal{H}_L(\mathbf{x}_L) \end{bmatrix}, \quad (1.130a)$$

$$\mathbf{H}_{K:L} = \begin{bmatrix} \mathbf{0} & \mathbf{H}_K & \mathbf{0} \\ \ddots & & \ddots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_L \end{bmatrix}, \quad (1.130b)$$

$$\mathbf{R}_{K:L} = \begin{bmatrix} \mathbf{R}_K & \mathbf{0} \\ \ddots & \\ \mathbf{0} & \mathbf{R}_L \end{bmatrix}. \quad (1.130c)$$

On obtient ainsi :

$$\mathbf{x}_{0:L}^a = \mathbf{x}_{0:L}^b + \mathbf{C}_{0:L}^a \mathbf{H}_{K:L}^T \mathbf{R}_{K:L}^{-1} (\mathbf{y}_{K:L} - \mathcal{H}_{K:L}(\mathbf{x}_{0:L}^b)), \quad (1.131a)$$

$$\mathbf{C}_{0:L}^a = (\mathbf{B}_{0:L}^{-1} + \mathbf{H}_{K:L}^T \mathbf{R}_{K:L}^{-1} \mathbf{H}_{K:L})^{-1}, \quad (1.131b)$$

Propagation : elle consiste à calculer la moyenne $\mathbf{x}_{S:L+S}^b$ et la covariance $\mathbf{B}_{S:L+S}$ de $p(\mathbf{x}_{S:L+S}|\mathbf{y}_{L:})$ à partir de la moyenne $\mathbf{x}_{0:L}^a$ et de la covariance $\mathbf{C}_{0:L}^a$ de $p(\mathbf{x}_{0:L}|\mathbf{y}_{L:})$. On met pour cela les équations (1.102b) sous la forme :

$$\mathbf{x}_{s:L+s} = \begin{bmatrix} \mathbf{x}_s^a \\ \vdots \\ \mathbf{x}_{L+s-1}^a \\ \mathcal{M}_{L+s}(\mathbf{x}_{L+s-1}^a) \end{bmatrix} + \begin{bmatrix} \mathbf{I}_n & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{L+s} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_s - \mathbf{x}_s^a \\ \vdots \\ \mathbf{x}_{L+s-1} - \mathbf{x}_{L+s-1}^a \\ \boldsymbol{\varepsilon}_{L+s}^q \end{bmatrix}, \quad (1.132)$$

pour $s = 1, \dots, S$. Ce qui nous permet de déduire le caractère gaussien de $p(\mathbf{x}_{s:L+s}|\mathbf{y}_{L:})$ et des relations de récurrence pour calculer sa moyenne et sa matrice de covariance :

$$\mathbf{x}_{s:L+s}^a = \begin{bmatrix} \mathbf{x}_{s:L+s-1}^a \\ \mathcal{M}_{L+s}(\mathbf{x}_{L+s-1}^a) \end{bmatrix}, \quad (1.133a)$$

$$\mathbf{C}_{0:L+s}^a = \begin{bmatrix} \mathbf{I}_n & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{L+s} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{C}_{s:L+s-1}^a & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{L+s} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{L+s} & \mathbf{I}_n \end{bmatrix}^T, \quad (1.133b)$$

pour $s = 1, \dots, S$. Finalement :

$$\mathbf{x}_{S:L+S}^b = \mathbf{x}_{S:L+S}^a, \quad (1.134a)$$

$$\mathbf{B}_{S:L+S} = \mathbf{C}_{S:L+S}^a, \quad (1.134b)$$

avec $\mathbf{C}_{S:L+S}^a$, le bloc diagonal de $\mathbf{C}_{0:L+S}^a$, correspondant aux indices temporels de S à $L + S$.

1.2.4.1 Remarque sur la méthode aller-retour

Un lisseur de Kalman $L = K = 1$ peut être obtenu à partir des résultats du filtre de Kalman $0 = K = L$ en ne marginalisant pas les densités à la fin de l'étape de propagation et en ajoutant une nouvelle étape d'analyse ([Anderson and Moore, 1979](#); [Cohn et al., 1994](#)) :

$$p(\mathbf{x}_0|\mathbf{y}_{:-1}) \xrightarrow{A} p(\mathbf{x}_0|\mathbf{y}_{:0}) \xrightarrow{P} p(\mathbf{x}_{0:1}|\mathbf{y}_{:0}) \xrightarrow{M} p(\mathbf{x}_1|\mathbf{y}_{:0}) \xrightarrow{A} \dots [\text{Filtre}] \quad (1.135a)$$

$$\xrightarrow{A} p(\mathbf{x}_{0:1}|\mathbf{y}_{:1}) \quad [\text{Lisseur}] \quad (1.135b)$$

En effet, à la fin de l'étape d'analyse, on dispose de la moyenne \mathbf{x}_0^a et covariance \mathbf{C}_0^a de $p(\mathbf{x}_0|\mathbf{y}_{:0})$. Puisque

$$\mathbf{x}_{0:1} = \begin{bmatrix} \mathbf{x}_0^a \\ \mathcal{M}_1(\mathbf{x}_0^a) \end{bmatrix} + \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{M}_1 & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 - \mathbf{x}_0^a \\ \boldsymbol{\varepsilon}_1^q \end{bmatrix}, \quad (1.136)$$

la densité $p(\mathbf{x}_{0:1}|\mathbf{y}_{:0})$ est gaussienne de moyenne et covariance :

$$\mathbf{x}_{0:1}^a = \begin{bmatrix} \mathbf{x}_0^a \\ \mathcal{M}_1(\mathbf{x}_0^a) \end{bmatrix}, \quad (1.137a)$$

$$\mathbf{C}_{0:1}^a = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{M}_1 & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{C}_0^a & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{M}_1 & \mathbf{I}_n \end{bmatrix}^T. \quad (1.137b)$$

La fonction de coût associée à la nouvelle densité a posteriori $p(\mathbf{x}_{0:1}|\mathbf{y}_{:1})$ est alors :

$$G(\mathbf{x}_{0:1}|\mathbf{y}_{:1}) \propto \frac{1}{2} \|\mathbf{y}_1 - \mathcal{H}_1(\mathbf{x}_1)\|_{\mathbf{R}_1}^2 + \frac{1}{2} \|\mathbf{x}_{0:1} - \mathbf{x}_{0:1}^a\|_{(\mathbf{C}_{0:1}^a)}^2. \quad (1.138)$$

Moyenne et matrice de covariance de $p(\mathbf{x}_{0:1}|\mathbf{y}_{:1})$ se calculent alors comme pour Eq. (1.128).

1.2.5 Le filtre de Kalman d'ensemble stochastique

Le filtre de Kalman d'ensemble stochastique ou EnKF (Burgers et al., 1998; Evensen, 1994, 2009; Houtekamer and Mitchell, 1998) est d'abord un filtre de Kalman de rang réduit permettant de travailler avec une fraction de la matrice de covariance. C'est ensuite une méthode permettant d'approcher séquentiellement la solution bayésienne du problème de filtrage sans reposer sur la linéarité des opérateurs. Dans la fenêtre d'assimilation $0 = K = L$, on suivra ici approximativement la succession d'étapes bayésiennes :

$$p(\mathbf{x}_0|\mathbf{y}_{:-1}) \xrightarrow{A} p(\mathbf{x}_0|\mathbf{y}_{:0}) \xrightarrow{P} p(\mathbf{x}_1|\mathbf{y}_{:0}). \quad (1.139)$$

Comme le filtre particulaire, l'EnKF utilise des ensembles. Cependant, ils sont utilisés pour estimer des moyennes et des matrices de covariance seulement.

Analyse : elle consiste à engendrer un ensemble a posteriori stochastique \mathbf{E}_0^a échantillonnant (jusqu'aux moments d'ordre deux) $p(\mathbf{x}_0|\mathbf{y}_{:0})$ à partir d'un ensemble stochastique \mathbf{E}_0^b échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:-1})$. Les poids associés à \mathbf{E}_0^b sont supposés ici identiques à $\frac{1}{M}$. Pour cela, une approche gaussienne de la densité jointe permet, comme dans la section 1.1.7, d'exprimer moyenne et matrice de covariance a posteriori en fonction de moments a priori. Par analogie avec la section 1.1.7, le meilleur estimateur de \mathbf{x}_0 fonction affine de \mathbf{y}_0 est tel que :

$$\mathbf{x}_0 = \tilde{\mathcal{K}}_0(\mathbf{y}_0) + \tilde{\boldsymbol{\eta}}_0, \quad (1.140)$$

avec $\tilde{\boldsymbol{\eta}}_0$ moyenne nulle et décorrélée de \mathbf{y}_0 :

$$\mathbb{E}_{p(\mathbf{x}_0, \mathbf{y}_0|\mathbf{y}_{:-1})}[\tilde{\boldsymbol{\eta}}_0] = \mathbf{0}_n, \quad (1.141a)$$

$$\mathbb{C}_{p(\mathbf{x}_0, \mathbf{y}_0|\mathbf{y}_{:-1})}[\mathbf{y}_0, \tilde{\boldsymbol{\eta}}_0] = \mathbf{0}_{d \times n}. \quad (1.141b)$$

Comme pour Eq. (1.78b, 1.78d), les moments a posteriori sont ensuite approchés par :

$$\mathbb{E}_{p(\mathbf{x}_0|\mathbf{y}_{:0})}[\mathbf{x}_0] \simeq \tilde{\mathcal{K}}_0(\mathbf{y}_0), \quad (1.142a)$$

$$\mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{:0})}[\mathbf{x}_0] \simeq \mathbb{C}_{p(\tilde{\boldsymbol{\eta}}_0|\mathbf{y}_{:-1})}[\tilde{\boldsymbol{\eta}}_0]. \quad (1.142b)$$

L'objectif de l'analyse est donc de produire des membres $(\mathbf{x}_0^{i,a})_{i=1, \dots, M}$ identiquement distribués à $p(\mathbf{x}_0|\mathbf{y}_{:0})$. Puisqu'on dispose seulement des approximations Eq. (1.142), on

se contentera de :

$$\mathbb{E}_{p(\mathbf{x}_0^{i,a}|\mathbf{y}_0)} [\mathbf{x}_0^{i,a}] = \tilde{\mathcal{K}}_0(\mathbf{y}_0), \quad (1.143a)$$

$$\mathbb{C}_{p(\mathbf{x}_0^{i,a}|\mathbf{y}_0)} [\mathbf{x}_0^{i,a}] = \mathbb{C}_{p(\tilde{\boldsymbol{\eta}}_0|\mathbf{y}_{:-1})} [\tilde{\boldsymbol{\eta}}_0], \quad (1.143b)$$

les moments d'ordre supérieurs à deux sont donc ignorés. A l'aide de l'ensemble a priori \mathbf{E}_0^b et d'un échantillon $(\boldsymbol{\varepsilon}_0^{i,r})_{i=1,\dots,M}$ de $p(\boldsymbol{\varepsilon}_0^r) = \mathcal{N}(\boldsymbol{\varepsilon}_0^r|\mathbf{0}_d, \mathbf{R}_0)$, on compose un ensemble $(\mathbf{x}_0^{i,b}, \boldsymbol{\varepsilon}_0^{i,r})_{i=1,\dots,M}$, conditionnellement indépendant de $(\mathbf{x}_0, \boldsymbol{\varepsilon}_0^r)$ sachant $\mathbf{y}_{:-1}$ et identiquement distribué selon $p(\mathbf{x}_0|\mathbf{y}_{:-1})p(\boldsymbol{\varepsilon}_0^r)$. Soit

$$\mathbf{x}_0^{i,a} = \tilde{\mathcal{K}}_0(\mathbf{y}_0) + \tilde{\boldsymbol{\eta}}_0^{i,b}, \quad (1.144a)$$

$$\tilde{\boldsymbol{\eta}}_0^{i,b} = \mathbf{x}_0^{i,b} - \tilde{\mathcal{K}}_0(\mathcal{H}(\mathbf{x}_0^{i,b}) + \boldsymbol{\varepsilon}_0^{i,r}), \quad (1.144b)$$

pour $i = 1, \dots, M$. Alors $\tilde{\boldsymbol{\eta}}_0^{i,b}$ est conditionnellement indépendant de $\mathbf{y}_0 = \mathcal{H}(\mathbf{x}_0) + \boldsymbol{\varepsilon}_0^r$ sachant $\mathbf{y}_{:-1}$: donc :

$$\mathbb{E}_{p(\mathbf{x}_0^{i,a}|\mathbf{y}_0)} [\mathbf{x}_0^{i,a}] = \tilde{\mathcal{K}}_0(\mathbf{y}_0) + \mathbb{E}_{p(\mathbf{x}_0^{i,b}, \boldsymbol{\varepsilon}_0^{i,r}|\mathbf{y}_0)} [\tilde{\boldsymbol{\eta}}_0^{i,b}], \quad (1.145a)$$

$$= \tilde{\mathcal{K}}_0(\mathbf{y}_0) + \mathbb{E}_{p(\mathbf{x}_0, \boldsymbol{\varepsilon}_0^r|\mathbf{y}_{:-1})} [\tilde{\boldsymbol{\eta}}_0], \quad (1.145b)$$

$$= \tilde{\mathcal{K}}_0(\mathbf{y}_0) \quad (1.145c)$$

$$\mathbb{C}_{p(\mathbf{x}_0^{i,a}|\mathbf{y}_0)} [\mathbf{x}_0^{i,a}] = \mathbb{C}_{p(\mathbf{x}_0^{i,b}, \boldsymbol{\varepsilon}_0^{i,r}|\mathbf{y}_0)} [\tilde{\boldsymbol{\eta}}_0^{i,b}], \quad (1.145d)$$

$$= \mathbb{C}_{p(\mathbf{x}_0, \boldsymbol{\varepsilon}_0^r|\mathbf{y}_{:-1})} [\tilde{\boldsymbol{\eta}}_0], \quad (1.145e)$$

$$= \mathbb{C}_{p(\tilde{\boldsymbol{\eta}}_0|\mathbf{y}_{:-1})} [\tilde{\boldsymbol{\eta}}_0], \quad (1.145f)$$

pour $i = 1, \dots, M$. L'ensemble a posteriori est alors $\mathbf{E}_0^a = [\mathbf{x}_0^{1,a} \ \dots \ \mathbf{x}_0^{M,a}]$.

Le paragraphe précédent suppose d'avoir $\tilde{\mathcal{K}}$ a disposition. Voyons à présent comment estimer une telle fonction affine. Puisque

$$\mathbf{x}_0^{i,a} = \mathbf{x}_0^{i,b} + \tilde{\mathbf{K}}_0(\mathbf{y}_0 - \mathcal{H}(\mathbf{x}_0^{i,b}) - \boldsymbol{\varepsilon}_0^{i,r}), \quad (1.146)$$

il suffit d'estimer la partie linéaire $\tilde{\mathbf{K}}_0$ de $\tilde{\mathcal{K}}_0$. Par analogie avec Eq. (1.78b), cette partie linéaire vérifie :

$$\tilde{\mathbf{K}}_0 = \mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{:-1})} [\mathbf{x}_0, \mathcal{H}_0(\mathbf{x}_0)] \left(\mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{:-1})} [\mathcal{H}_0(\mathbf{x}_0)] + \mathbf{R}_0 \right)^{-1}. \quad (1.147)$$

Les covariances de cette expression peuvent être estimées à l'aide de Eq. (1.27a) par :

$$\mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{:-1})}[\mathbf{x}_0, \mathcal{H}_0(\mathbf{x}_0)] \simeq \mathbf{E}_0^b \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{M-1} \right) \mathcal{H}(\mathbf{E}_0^b)^T, \quad (1.148a)$$

$$\mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{:-1})}[\mathcal{H}_0(\mathbf{x}_0)] \simeq \mathcal{H}(\mathbf{E}_0^b) \left(\frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}}{M-1} \right) \mathcal{H}(\mathbf{E}_0^b)^T. \quad (1.148b)$$

Propagation : elle consiste à engendrer un ensemble \mathbf{E}_1^a échantillonnant $p(\mathbf{x}_1|\mathbf{y}_{:0})$ à partir de l'ensemble \mathbf{E}_0^a échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:0})$. De même que la propagation du filtre particulaire Eq. (1.119) on utilisera :

$$\mathbf{x}_1^{i,a} = \mathcal{M}_1(\mathbf{x}_0^{i,a}) + \boldsymbol{\varepsilon}_1^{i,q}, \quad (1.149)$$

avec $(\boldsymbol{\varepsilon}_1^{i,q})_{i=1,\dots,M}$ un ensemble indépendant identiquement distribué selon la densité des erreurs modèles $\mathcal{N}(\boldsymbol{\varepsilon}_1^q|\mathbf{0}_n, \mathbf{Q}_1)$.

En pratique, l'analyse du filtre de Kalman d'ensemble permet d'obtenir avec peu de membres un ensemble a posteriori dont la moyenne et la matrice de covariance sont une bonne approximation de la moyenne et de la covariance a posteriori. En revanche, les moments d'ordre supérieurs sont ignorés. Ainsi, même pour un nombre de membres infini, le filtre de Kalman d'ensemble ne résout pas exactement le problème de filtrage bayésien. Puisque l'EnKF ne nécessite que peu de membres pour être précis, il est performant en terme de mémoire. Un ensemble est en effet moins volumineux qu'une matrice de covariance lorsque $M \ll n$. La propagation ne requiert pas de produits matriciels et l'évaluation du modèle sur chaque membre de l'ensemble est facilement parallélisable.

Cependant, lorsque $M \ll n$ l'estimation des matrices de covariance introduit des erreurs d'échantillonnages qu'il faut corriger pour pouvoir tirer pleinement profit des capacités du filtre de Kalman d'ensemble.

1.2.5.1 Localisation et inflation

En géophysique, l'état représente des paramètres du système en chaque point du globe terrestre. On observe alors fréquemment une faible corrélation entre des points éloignés. Cette connaissance a priori peut être utilisée pour corriger l'estimation des matrices de covariance et réduire l'erreur d'échantillonnage. Deux principales méthodes sont employées pour rendre locale la matrice de covariance. La première est la localisation en domaine et consiste à faire plusieurs analyses indépendantes pour chaque domaine du globe. Pour chacune de ces analyses, seules les observations à proximité du domaine

sont utilisées. La seconde est la localisation par produit de Schur et consiste à multiplier terme à terme la matrice estimée avec une seconde matrice atténuant les corrélations à longue distance. Pour un traitement plus complet de la localisation on pourra consulter [Houtekamer and Mitchell \(2001\)](#), [Hamill et al. \(2001\)](#), [Evensen \(2003\)](#), [Ott et al. \(2004\)](#).

Pour estimer l'état \mathbf{x} , on utilise une estimation de la moyenne et de la covariance a priori. L'estimation étant inexacte, il en résulte une incertitude sur ces moments. Cependant, l'algorithme n'en tient pas compte. Il sous-estime alors l'incertitude sur \mathbf{x} . Puisque la variance a posteriori est une mesure de l'incertitude sur \mathbf{x} , l'algorithme la sous-estime aussi. Lors du prochain cycle, l'incertitude sur les moments a priori sera alors d'autant plus importante. On observe donc une amplification de l'incertitude à chaque cycle qui mène nécessairement à une divergence de l'algorithme si elle n'est pas compensée. C'est le rôle de l'inflation. Elle consiste à dilater les matrices de covariances par un facteur multiplicatif afin de corriger l'erreur d'échantillonnage et de mieux estimer l'incertitude sur \mathbf{x} .

1.2.6 Le filtre de Kalman d'ensemble déterministe

On peut reprocher au filtre de Kalman d'ensemble stochastique d'ajouter des erreurs d'échantillonnage lors de l'introduction des erreurs d'observations $\boldsymbol{\varepsilon}_0^{i,r}$ dans Eq. (1.146). L'alternative proposée dans [Bishop et al. \(2001\)](#) et [Hunt et al. \(2007\)](#) est d'engendrer l'ensemble a posteriori de manière déterministe plutôt que stochastique.

Analyse : elle consiste à engendrer un ensemble a posteriori déterministe \mathbf{E}_0^a échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:0})$ à partir d'un ensemble déterministe \mathbf{E}_0^b échantillonnant $p(\mathbf{x}_0|\mathbf{y}_{:-1})$. Soit

$$\begin{bmatrix} \mathbf{x}_0^b & \mathbf{X}_0^b \end{bmatrix} = \mathbf{E}_0^b \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{M-1}} \end{bmatrix}, \quad (1.150)$$

la moyenne empirique et une matrice de déviation de \mathbf{E}_0^b calculées avec une base orthonormale $\begin{bmatrix} \frac{\mathbf{1}_M}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}$ (cf 1.1.5). Quitte à supprimer des membres redondants de \mathbf{E}_0^b , on supposera que cet ensemble est injectif dans le sens où la matrice associée est injective. Dans ce cas, \mathbf{X}_0^b est injective. Par projection orthogonale de $\mathbf{x}_0 - \mathbf{x}_0^b$ sur $\text{Im}(\mathbf{X}_0^b) \oplus \text{Im}(\mathbf{X}_0^b)^\perp$, cette injectivité assure l'existence et l'unicité d'un couple aléatoire $(\mathbf{w}, \boldsymbol{\varepsilon}^e)$ à valeurs dans $\mathbb{R}^m \times \mathbb{R}^n$, $m = M - 1$ tel que :

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w} + \boldsymbol{\varepsilon}^e, \quad (1.151)$$

on nommera \mathbf{w} variable d'ensemble et $\boldsymbol{\varepsilon}^e$ erreur d'échantillonnage. En effet, si l'estima-

tion empirique des moments a priori est exacte :

$$\mathbb{E}_{p(\mathbf{x}_0|\mathbf{y}_{-1})}[\mathbf{x}_0] = \mathbf{x}_0^b, \quad (1.152a)$$

$$\mathbb{C}_{p(\mathbf{x}_0|\mathbf{y}_{-1})}[\mathbf{x}_0] = (\mathbf{X}_0^b) (\mathbf{X}_0^b)^\top, \quad (1.152b)$$

alors $\boldsymbol{\varepsilon}^e = \mathbf{0}_n$ et la variable d'ensemble est centrée et réduite :

$$\mathbb{E}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathbf{w}] = \mathbf{0}_m, \quad (1.153a)$$

$$\mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathbf{w}] = \mathbf{I}_m, \quad (1.153b)$$

d'après l'annexe A.4. Un ensemble déterministe échantillonnant $p(\mathbf{w}|\mathbf{y}_{-1})$ est donc :

$$\mathbf{D}^b = \sqrt{m} \mathbf{U}_M^\top. \quad (1.154)$$

On peut appliquer l'analyse Eq. (1.85) de la section 1.1.7 à la variable d'ensemble \mathbf{w} avec l'opérateur d'observation $\mathcal{F}_0(\mathbf{w}) = \mathcal{H}_0(\mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w})$. On obtient les approximations suivantes des moments a posteriori :

$$\mathbb{E}_{p(\mathbf{w}|\mathbf{y}_0)}[\mathbf{w}] \simeq \mathbf{w}^a, \quad (1.155a)$$

$$\mathbf{w}^a = (\tilde{\mathbf{F}}_0^\top \tilde{\mathbf{R}}_0^{-1} \tilde{\mathbf{F}}_0 + \mathbf{I}_m)^{-1} \tilde{\mathbf{F}}_0^\top \tilde{\mathbf{R}}_0^{-1} (\mathbf{y}_0 - \tilde{\mathbf{f}}_0), \quad (1.155b)$$

$$\mathbb{C}_{p(\mathbf{w}|\mathbf{y}_0)}[\mathbf{w}] \simeq (\mathbf{W}^a) (\mathbf{W}^a)^\top, \quad (1.155c)$$

$$(\mathbf{W}^a) (\mathbf{W}^a)^\top = (\tilde{\mathbf{F}}_0^\top \tilde{\mathbf{R}}_0^{-1} \tilde{\mathbf{F}}_0 + \mathbf{I}_m)^{-1}, \quad (1.155d)$$

avec

$$\tilde{\mathbf{f}}_0 = \mathbb{E}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathcal{F}(\mathbf{w})], \quad (1.156a)$$

$$\tilde{\mathbf{F}}_0 = \mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathcal{F}(\mathbf{w}), \mathbf{w}] \mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathbf{w}]^{-1}, \quad (1.156b)$$

$$= \mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathcal{F}(\mathbf{w}), \mathbf{w}], \quad (1.156c)$$

$$\tilde{\mathbf{R}} = \mathbf{R} + \mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathcal{F}(\mathbf{w})] - (\tilde{\mathbf{F}}_0) (\tilde{\mathbf{F}}_0)^\top. \quad (1.156d)$$

Notons que l'inversion se passe ici dans l'espace de l'ensemble dont la dimension m est généralement bien plus petite que la dimension de l'espace des états n ou des observations d . Les moments de ces expressions peuvent être estimés à l'aide de l'ensemble Eq. (1.154) :

$$\tilde{\mathbf{f}}_0 \simeq \mathcal{F}(\mathbf{D}^b) \frac{\mathbf{1}_M}{M}, \quad (1.157a)$$

$$\tilde{\mathbf{F}}_0 \simeq \mathcal{F}(\mathbf{D}^b) \left(\frac{\mathbf{U}_M \mathbf{U}_M^T}{m} \right) (\mathbf{D}^b)^T, \quad (1.157b)$$

$$\simeq \mathcal{F}(\mathbf{D}^b) \frac{1}{\sqrt{m}} \mathbf{U}_M \left(\sqrt{m} \mathbf{U}_M^T \frac{1}{\sqrt{m}} \mathbf{U}_M \right)^T, \quad (1.157c)$$

$$\simeq \mathcal{F}(\mathbf{D}^b) \frac{1}{\sqrt{m}} \mathbf{U}_M \quad (1.157d)$$

$$\mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{-1})}[\mathcal{F}(\mathbf{w})] \simeq \mathcal{F}(\mathbf{D}^b) \left(\frac{\mathbf{U}_M \mathbf{U}_M^T}{m} \right) \mathcal{F}(\mathbf{D}^b)^T, \quad (1.157e)$$

$$\simeq \tilde{\mathbf{F}}_0 \tilde{\mathbf{F}}_0^T, \quad (1.157f)$$

et se réduisent à :

$$\begin{bmatrix} \tilde{\mathbf{f}}_0 & \tilde{\mathbf{F}}_0 \end{bmatrix} \simeq \mathcal{F}(\mathbf{D}^b) \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{1}{\sqrt{m}} \mathbf{U}_M \end{bmatrix}, \quad (1.158a)$$

$$\tilde{\mathbf{R}} \simeq \mathbf{R}. \quad (1.158b)$$

Puisque $\tilde{\mathbf{R}} \simeq \mathbf{R}$, on obtient, d'après les résultats de la section 1.1.7, les mêmes approximations des moments a posteriori si, lors de l'analyse de \mathbf{w} , on avait seulement remplacé \mathcal{F} par sa meilleure approximation affine $\tilde{\mathcal{F}}_0(\mathbf{w}) = \tilde{\mathbf{f}}_0 + \tilde{\mathbf{F}}_0 \mathbf{w}$ au sens des erreurs quadratiques moyennes selon $p(\mathbf{w}|\mathbf{y}_{-1})$.

Finalement, un ensemble déterministe a posteriori \mathbf{D}^a ayant \mathbf{w}^a comme moyenne empirique et $(\mathbf{W}^a)(\mathbf{W}^a)^T$ comme covariance empirique est engendré avec :

$$\mathbf{D}^a = \begin{bmatrix} \mathbf{w}^a & \mathbf{W}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T, \quad (1.159)$$

d'après la section 1.1.5. En appliquant la transformation :

$$\mathbf{E}_0^a = \mathbf{x}_0^b \mathbf{1}_M^T + \mathbf{X}_0^b \mathbf{D}^a, \quad (1.160)$$

on obtient un ensemble a posteriori déterministe \mathbf{E}_0^a échantillonnant (approximativement) $p(\mathbf{x}_0|\mathbf{y}_{:0})$.

Propagation : elle est identique à la propagation du filtre de Kalman d'ensemble stochastique.

1.2.7 Méthode variationnelle : 4DVar

La méthode 4DVar (Dimet and Talagrand, 1986) généralise la méthode variationnelle présentée section 1.1.8 à une fonction de coût s'étendant sur plusieurs pas de temps. L'avantage des méthodes d'optimisation numérique est de pouvoir gérer la non linéarité des opérateurs. Dans la fenêtre d'assimilation $0 \leq K \leq L$, on suivra ici approximativement la succession bayésienne d'étapes :

$$p(\mathbf{x}_{0:L}) \xrightarrow{A} p(\mathbf{x}_{0:L}|\mathbf{y}_{0:L}) \xrightarrow{P} p(\mathbf{x}_{0:L+S}|\mathbf{y}_{0:L}). \quad (1.161)$$

L'étape de propagation étant impossible dans le cadre de l'estimation par maximum a posteriori, le caractère séquentiel du 4DVar s'arrête ici. C'est pourquoi les observations antérieures à t_0 ne sont pas ici prises en compte.

Analyse : Elle consiste à calculer le maximiseur de $p(\mathbf{x}_{0:L}|\mathbf{y}_{0:L})$. La densité a priori $p(\mathbf{x}_{0:L})$ se décompose de la façon suivante :

$$p(\mathbf{x}_{0:L}) = p(\mathbf{x}_0) \prod_{l=1}^L p(\boldsymbol{\varepsilon}_l^q = \mathbf{x}_l - \mathcal{M}_l(\mathbf{x}_{l-1})), \quad (1.162)$$

grâce à l'indépendance des erreurs modèle. Avec l'hypothèse $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0|\mathbf{x}_0^b, \mathbf{B}_0)$, la fonction de coût a posteriori Eq. (1.111) vérifie alors :

$$\begin{aligned} G(\mathbf{x}_{0:L}|\mathbf{y}_{0:L}) &\propto \sum_{l=0}^L \frac{1}{2} \|\mathbf{y}_l - \mathcal{H}_l(\mathbf{x}_l)\|_{\mathbf{R}_l^{-1}}^2 + \sum_{l=1}^L \frac{1}{2} \|\mathbf{x}_l - \mathcal{M}_l(\mathbf{x}_{l-1})\|_{\mathbf{Q}_l^{-1}}^2 \\ &\quad + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0^{-1}}^2. \end{aligned} \quad (1.163)$$

L'analyse variationnelle consiste à minimiser itérativement cette fonction de coût comme dans la section 1.1.8, sauf que les linéarisations concernent maintenant les opérateurs d'observation et le modèle $(\mathcal{H}_l, \mathcal{M}_l)_{l=K, \dots, L}$. Au terme de cette analyse, on dispose du maximum a posteriori :

$$\mathbf{x}_{0:L}^* = \underset{\mathbf{x}_{0:L}}{\operatorname{argmin}} G(\mathbf{x}_{0:L}|\mathbf{y}_{0:L}). \quad (1.164)$$

Cette méthode n'est cependant pas directement applicable en grande dimension car la minimisation selon la trajectoire $\mathbf{x}_{0:L}$ multiplie la dimension de l'espace des états par la longueur de la fenêtre d'assimilation. C'est pourquoi la méthode du 4DVar est généralement utilisée dans un cadre fortement contraint où les erreurs modèle $\{\boldsymbol{\varepsilon}_l^q\}_{l=1, \dots, L}$

sont ignorées. Ainsi, la trajectoire $\mathbf{x}_{0:L}$ est déterminée par son état initial \mathbf{x}_0 :

$$\mathbf{x}_l = \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0). \quad (1.165)$$

pour $l = 1, \dots, L$. Ce qui a pour effet, d'après Eq. (1.114), de supprimer tous les termes liés à l'erreur modèle dans Eq. (1.163). La minimisation ne se fait alors plus que selon \mathbf{x}_0 :

$$G(\mathbf{x}_0 | \mathbf{y}_{0:L}) \propto \sum_{l=0}^L \frac{1}{2} \|\mathbf{y}_l - \mathcal{H}_l \circ \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0)\|_{\mathbf{R}_l}^2 + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0}^2. \quad (1.166)$$

Propagation : elle pourrait consister à calculer le minimiseur de $G(\mathbf{x}_{0:L+S} | \mathbf{y}_{0:L})$ lors d'une phase correspondant à Eq. (1.107b). De même que pour Eq. (1.162) l'indépendance des erreurs permet de décomposer cette fonction de coût en :

$$G(\mathbf{x}_{0:L+S} | \mathbf{y}_{0:L}) \propto \sum_{l=L+1}^{L+S} \frac{1}{2} \|\mathbf{x}_{l+S} - \mathcal{M}_{l+S}(\mathbf{x}_{l+S-1})\|_{\mathbf{Q}_l}^2 - G(\mathbf{x}_{0:L} | \mathbf{y}_{0:L}). \quad (1.167)$$

Avec l'analyse précédente, la solution de ce problème est triviale, le minimiseur recherché est :

$$\mathbf{x}_{0:L}^* = \mathbf{x}_{0:L}^*, \quad (1.168a)$$

$$\mathbf{x}_l^* = \mathcal{M}_{l \leftarrow L}(\mathbf{x}_L^*), \quad (1.168b)$$

$$l = L+1, \dots, L+S. \quad (1.168c)$$

Dans une seconde phase correspondant à Eq. (1.108), il faudrait calculer le maximiseur de :

$$p(\mathbf{x}_{S:L+S} | \mathbf{y}_{0:L}) = \int p(\mathbf{x}_{0:L+S} | \mathbf{y}_{0:L}) d\mathbf{x}_{0:S-1}. \quad (1.169)$$

Cette intégration marque la fin du processus de cyclage car, sans plus de développements extérieurs au cadre du 4DVar, on ne sait ni calculer, ni approcher ce maximiseur.

1.2.8 Conclusion de l'assimilation de données avec dimension temporelle

La règle de Bayes nous permet de calculer séquentiellement la densité a posteriori. Chaque cycle se décompose alors en une étape d'analyse qui calcule la densité a posteriori à partir de la densité a priori et une étape de propagation qui calcule la densité a priori du prochain cycle en translatant la fenêtre d'assimilation.

Le filtre particulaire utilise une représentation ensembliste des densités. L'analyse consiste en un échantillonnage d'importance transformant l'ensemble a priori en ensemble a posteriori. La propagation consiste en une application terme à terme du modèle à l'ensemble a posteriori pour obtenir l'ensemble a priori du prochain cycle. Le filtre particulaire est simple et requiert peu d'hypothèses mais il demande beaucoup d'évaluations d'opérateurs en grande dimension.

Le filtre de Kalman utilise une représentation gaussienne des densités. Des hypothèses de linéarité sont faites pour conserver ce caractère gaussien. L'analyse consiste alors à transformer algébriquement la moyenne et la matrice de covariance a priori en la moyenne et la matrice de covariance a posteriori. La propagation utilise la linéarité du modèle pour transformer la moyenne et la matrice de covariance a posteriori en la moyenne et la matrice de covariance a priori du prochain cycle.

Le filtre de Kalman d'ensemble utilise des représentations ensemblistes et gaussiennes des densités. Lors de l'analyse, la moyenne et la matrice de covariance a priori sont estimées à partir de l'ensemble à priori. Elles engendrent une approximation gaussienne de la densité jointe qui nous permet d'approcher la moyenne et la covariance a posteriori. Pour engendrer un ensemble a posteriori qui conserve ces moments, il suffit d'estimer le gain de Kalman à l'aide de l'ensemble a priori puis de l'appliquer à l'ensemble a priori avec des perturbations de l'observation. La propagation est la même que pour le filtre particulaire.

Le filtre de Kalman d'ensemble déterministe est semblable au filtre de Kalman d'ensemble stochastique. La différence est que l'ensemble a posteriori est engendré de manière à ce que sa moyenne et sa matrice de covariance empiriques soient égales à l'approximation de la moyenne et de la covariance produites par l'analyse.

Le 4DVar utilise la représentation analytique de la densité a posteriori donnée par la règle de Bayes. L'analyse consiste à calculer le maximum a posteriori en minimisant itérativement la fonction de coût a posteriori. Bien que cette fonction de coût s'étende sur plusieurs pas de temps, la méthode n'est pas pleinement séquentielle. En effet, on ne peut calculer, sans d'autres résultats extérieurs au cadre du 4DVar, d'estimateur de second ordre permettant d'engendrer un ensemble a posteriori.

Dans la section suivante, nous allons voir des solutions permettant de bénéficier de la précision des méthodes variationnelles dans un cadre séquentiel.

1.3 L'assimilation de données ensemble variationnelle fortement contrainte

Dans un cadre séquentiel, les méthodes variationnelles présentent un avantage quant à l'estimation de l'état d'un système à la dynamique non linéaire grâce aux techniques d'optimisation numériques. En revanche, l'estimation des covariances n'est pas naturelle. Inversement, les méthodes ensemblistes sont numériquement efficaces pour estimer ces covariances séquentiellement mais leur analyse n'est pas itérative. Les méthodes ensemble variationnelles (EnVar) tentent de combiner ces deux approches pour profiter de leurs atouts respectifs. Dans cette section, nous revoyons les méthodes EnVar présentes dans la littérature dans un cadre fortement contraint, c'est à dire sans erreurs modèle. On suivra pour cela le chapitre 7 de [Asch et al. \(2017\)](#).

1.3.1 Méthodes hybrides

L'idée de [Hamill and Snyder \(2000\)](#) est d'engendrer une matrice de covariance d'ébauche \mathbf{B}_0 dépendante du temps par combinaison convexe d'une matrice de covariance empirique et de la matrice de covariance climatologique.

Si $\mathbf{E}_0^b = [\mathbf{x}_0^{1,b}, \dots, \mathbf{x}_0^{M,b}]$ est l'ensemble a priori, \mathbf{A}_0^b son anomalie normalisée et \mathbf{B} une matrice de covariance climatologique indépendante du temps, cette combinaison est alors :

$$\mathbf{B}_0 = \alpha \mathbf{B} + (1 - \alpha) (\mathbf{A}_0^b) (\mathbf{A}_0^b)^T, \quad (1.170)$$

avec $\alpha \in [0, 1]$ un coefficient à régler. Ainsi, la matrice \mathbf{B}_0 est inversible même lorsque $(\mathbf{A}_0^b) (\mathbf{A}_0^b)^T$ ne l'est pas. L'ensemble a posteriori $\mathbf{E}_0^a = [\mathbf{x}_0^{1,a}, \dots, \mathbf{x}_0^{M,a}]$ est ensuite engendré par M analyses variationnelles parallèles :

$$\mathbf{x}_0^{i,a} = \operatorname{argmin}_{\mathbf{x}_0} \frac{1}{2} \left\| \mathbf{y}_0 + \boldsymbol{\varepsilon}_0^{i,r} - \mathcal{H}_0(\mathbf{x}_0) \right\|_{\mathbf{R}_0^{-1}}^2 + \frac{1}{2} \left\| \mathbf{x}_0 - \mathbf{x}_0^{i,b} \right\|_{\mathbf{B}_0^{-1}}^2, \quad (1.171)$$

avec $i = 1, \dots, M$ et $(\boldsymbol{\varepsilon}_0^{i,r})_{i=1, \dots, M}$ un ensemble indépendant et identiquement distribué à $\mathcal{N}(\boldsymbol{\varepsilon}_0^r | \mathbf{0}_d, \mathbf{R}_0)$. Pour clore le cycle d'assimilation l'ensemble a posteriori \mathbf{E}_0^a est alors propagé à l'aide du modèle pour engendrer l'ensemble a priori \mathbf{E}_1^b du prochain cycle.

Nous avons vu dans le cadre du filtre de Kalman d'ensemble stochastique que la perturbation des données permettait de générer un ensemble a posteriori ayant de bonnes statistiques jusqu'à l'ordre deux. Cette propriété reposait sur le fait que l'estimation de la moyenne a posteriori produite par le filtre de Kalman d'ensemble était une fonction affine de \mathbf{y}_0 . Puisque le minimiseur de la fonction de coût Eq. (1.171) n'est pas une fonction

affine de \mathbf{y}_0 , il n'est pas nécessaire que cette propriété se conserve. Cependant, [Jardak and Talagrand \(2018\)](#) montre numériquement que l'ensemble a posteriori ainsi engendré variationnellement a de bonnes statistiques. En revanche, la perturbation stochastique des observations introduit nécessairement des erreurs d'échantillonnage et les multiples analyses variationnelles, bien que parallélisables, restent coûteuses. Pour améliorer les performances de cette méthode, il est possible d'engendrer l'ensemble a posteriori de manière déterministe, c'est à dire sans perturber l'observation ([Auligné et al., 2016](#); [Bocquet et al., 2015](#); [Sakov and Bertino, 2011](#)). La localisation de la covariance estimée peut aussi être envisagée.

1.3.2 Ensemble d'assimilations

Il est possible d'étendre le principe précédent à une fenêtre d'assimilation plus générale. C'est à dire minimiser en parallèle de multiples fonctions de coût de type 4DVar.

On engendre donc un ensemble a posteriori $\mathbf{E}_0^a = [\mathbf{x}_0^{1,a}, \dots, \mathbf{x}_0^{M,a}]$ à partir d'un ensemble a priori $\mathbf{E}_0^b = [\mathbf{x}_0^{1,b}, \dots, \mathbf{x}_0^{M,b}]$ et d'une matrice de covariance a priori \mathbf{B}_0 . Cette matrice de covariance peut être une matrice climatologique ou la covariance empirique (inflatée et localisée) de l'ensemble a priori ([Raynaud et al., 2009](#)) ou une combinaison convexe des deux. L'ensemble a posteriori $\mathbf{E}_0^a = [\mathbf{x}_0^{1,a}, \dots, \mathbf{x}_0^{M,a}]$ est ensuite engendré par M analyses variationnelles parallèles :

$$\mathbf{x}_0^{i,a} = \operatorname{argmin}_{\mathbf{x}_0} \frac{1}{2} \sum_{l=0}^L \left\| \mathbf{y}_l + \boldsymbol{\varepsilon}_l^{i,r} - \mathcal{H}_l \circ \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0) \right\|_{\mathbf{R}_0^{-1}}^2 + \frac{1}{2} \left\| \mathbf{x}_0 - \mathbf{x}_0^{i,b} \right\|_{\mathbf{B}_0^{-1}}^2, \quad (1.172)$$

pour $i = 1, \dots, M$ et $(\boldsymbol{\varepsilon}_0^i)_{i=1, \dots, M}$ un ensemble mutuellement indépendant identiquement distribué à $\mathcal{N}(\boldsymbol{\varepsilon}_0 | \mathbf{0}_d, \mathbf{R}_0)$.

1.3.3 Méthodes 4D variationnelles d'ensemble

Les méthodes d'optimisation numériques misent en œuvre lors de l'analyse variationnelle reposent sur l'évaluation du gradient de la fonction de coût. Cette évaluation fait elle-même intervenir l'adjoint (ou la transposée de la matrice jacobienne) du modèle et de l'opérateur d'observation. Pour des opérateurs complexes, ce code est difficile à manipuler. L'idée des méthodes 4D variationnelles d'ensemble est d'estimer ces adjoints à l'aide des différences finies ([Liu et al., 2008](#)). Cependant, pour estimer la dérivée d'une fonction par différence finie, il faut pouvoir l'évaluer pour un petit déplacement dans chaque direction de l'espace. En grande dimension, cela est numériquement coûteux. Une alternative consiste alors à restreindre le problème d'optimisation variationnelle à

un sous-espace de l'espace des états totalisant une grande fraction de la variabilité a priori de \mathbf{x}_0 (Robert et al., 2006, 2005).

L'analyse d'une méthode 4DEnVar suppose un ensemble a priori $\mathbf{E}_0^b \in \mathbb{R}^{n \times M}$ échantillonnant $p(\mathbf{x}_0 | \mathbf{y}_{:L-S})$. Soit :

$$\begin{bmatrix} \mathbf{x}_0^b & \mathbf{X}_0^b \end{bmatrix} = \mathbf{E}_0^b \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{M-1}} \end{bmatrix}, \quad (1.173)$$

la moyenne empirique et une matrice de déviation injective de \mathbf{E}_0^b , respectivement. La covariance empirique a priori est alors :

$$\mathbf{B}_0 = (\mathbf{X}_0^b) (\mathbf{X}_0^b)^\top, \quad (1.174)$$

d'après la section 1.1.5. Si le nombre de membres M de l'ensemble est petit devant n , la covariance \mathbf{B}_0 n'est pas surjective, l'approximation gaussienne de la densité a priori suivante n'est donc pas bien définie :

$$p(\mathbf{x}_0 | \mathbf{y}_{:L-S}) \simeq c e^{-\frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{(\mathbf{B}_0)^{-1}}^2}, \quad (1.175)$$

avec c une constante normalisatrice. On ne dispose donc pas d'expression analytique de la fonction de coût a posteriori $G(\mathbf{x}_0 | \mathbf{y}_{:L-S})$. Cependant, si on néglige les erreurs d'échantillonnage, on obtient d'après Eq. (1.151) l'existence d'une variable d'ensemble \mathbf{w} centrée, réduite et à valeur dans \mathbb{R}^m , $m = M - 1$ telle que :

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w}. \quad (1.176)$$

En appliquant Eq. (1.109a) à cette variable d'ensemble, on obtient, par passage au logarithme, une expression de la fonction de coût a posteriori de \mathbf{w} :

$$G(\mathbf{w} | \mathbf{y}_{:L}) \propto \frac{1}{2} \sum_{l=K}^L \|\mathbf{y}_l - \mathcal{F}_l(\mathbf{w})\|_{\mathbf{R}_l^{-1}}^2 - \ln p(\mathbf{w} | \mathbf{y}_{:L-S}), \quad (1.177)$$

avec $\mathcal{F}_l(\mathbf{w}) = \mathcal{H}_l \circ \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w})$ pour $l = K, \dots, L$. À la différence de la variable d'état, la variable d'ensemble est centrée et réduite. Sa fonction de coût a priori admet donc une approximation quadratique bien définie :

$$-\ln p(\mathbf{w} | \mathbf{y}_{:L-S}) \simeq c + \frac{1}{2} \|\mathbf{w}\|^2, \quad (1.178)$$

avec c une constante normalisatrice. En l'injectant dans Eq. (1.177), on obtient la fonc-

tion de coût a posteriori dans l'espace de l'ensemble :

$$J(\mathbf{w}|\mathbf{y}_{:L}) = \frac{1}{2} \sum_{l=K}^L \|\mathbf{y}_l - \mathcal{F}_l(\mathbf{w})\|_{\mathbb{R}^1}^2 + \frac{1}{2} \|\mathbf{w}\|^2, \quad (1.179)$$

l'utilisation de J plutôt que G insiste sur le fait que $J(\mathbf{w}|\mathbf{y}_{:L})$ n'est qu'une approximation de $G(\mathbf{w}|\mathbf{y}_{:L}) \propto -\ln p(\mathbf{w}|\mathbf{y}_{:L})$.

Le changement de variable dans l'espace de l'ensemble a donc pour effet de réduire la dimension de la variable de contrôle de n à $m = M - 1$ et de préconditionner le terme a priori de la fonction de coût. La minimisation variationnelle de cette fonction de coût à l'aide de méthodes d'optimisation numérique requiert, comme dans la méthode de Gauss-Newton, l'évaluation de la dérivée (que l'on confondra avec la matrice jacobienne) des $(\mathcal{F}_l)_{l=K,\dots,L}$. Soit \mathbf{w}^j l'itéré courant de la méthode d'optimisation, les différences finies permettent d'estimer les dérivées de \mathcal{F}_l au point \mathbf{w}^j à l'aide de son évaluation sur un ensemble contracté et centré sur \mathbf{w}^j . Un tel ensemble s'obtient avec :

$$\mathbf{D}^j = \begin{bmatrix} \mathbf{w}^j & \epsilon \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T, \quad (1.180)$$

où $\epsilon \ll 1$ est le pas des différences finies et la matrice \mathbf{U}_M est telle que la base $\begin{bmatrix} \frac{1}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}$ soit orthonormale. Les différences finies s'écrivent alors :

$$\begin{bmatrix} \mathcal{F}_l(\mathbf{w}^j) & \frac{d\mathcal{F}_l}{d\mathbf{w}}(\mathbf{w}^j) \end{bmatrix} \simeq \mathcal{F}_l(\mathbf{D}^j) \begin{bmatrix} \frac{1}{M} & \frac{1}{\epsilon\sqrt{m}} \mathbf{U}_M \end{bmatrix}. \quad (1.181)$$

Pour terminer de spécifier l'analyse d'une méthode 4D-EnVar, il faut engendrer un ensemble a posteriori. Pour cela, on peut utiliser un ensemble d'assimilations (Buehner et al., 2010). Il est aussi possible d'engendrer l'ensemble a posteriori de manière déterministe comme dans le 4D-LETKF (Fertig et al., 2007; Hunt et al., 2004) ou le lisseur de Kalman itératif abordé dans la section suivante.

1.3.4 Le lisseur de Kalman itératif

Le lisseur de Kalman itératif (IEnKS, Bocquet and Sakov, 2014; Sakov et al., 2012) est une méthode d'assimilation de données 4D variationnelle d'ensemble approchant séquentiellement la solution bayésienne du problème de lissage. Dans la fenêtre d'assimilation $0 \leq K \leq L$ et en négligeant les erreurs modèle, on suivra ici approximativement la succession d'étapes bayésiennes :

$$p(\mathbf{x}_0|\mathbf{y}_{:L-S}) \xrightarrow{\text{A}} p(\mathbf{x}_0|\mathbf{y}_{:L}) \xrightarrow{\text{P}} p(\mathbf{x}_S|\mathbf{y}_{:L}). \quad (1.182)$$

L'IEnKS utilise une analyse variationnelle dans l'espace réduit de l'ensemble où les adjoints des opérateurs sont estimés à l'aide des différences finies. Nous présenterons ici l'IEnKS avec deux variantes (bundle et transform) de la méthode de Gauss-Newton. D'autres méthodes d'optimisation numériques plus élaborées sont envisageables, par exemple les méthodes de Quasi-Newton, de Levenberg-Marquardt ou de région de confiance (Bocquet and Sakov, 2012; Mandel et al., 2016). Les ensembles sont engendrés de manière déterministe, ce qui permet l'estimation de matrices de covariance au cours du temps.

Analyse : On se donne par hypothèse de récurrence un ensemble a priori $\mathbf{E}_0^b \in \mathbb{R}^{n \times M}$ échantillonnant $p(\mathbf{x}_0 | \mathbf{y}_{:L-S})$. À l'aide d'une base orthonormale $\begin{bmatrix} \frac{1}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}$, on calcule la moyenne empirique et une matrice de déviation injective via :

$$\begin{bmatrix} \mathbf{x}_0^b & \mathbf{X}_0^b \end{bmatrix} = \mathbf{E}_0^b \begin{bmatrix} \frac{1}{M} & \frac{1}{\sqrt{M-1}} \mathbf{U}_M \end{bmatrix}. \quad (1.183)$$

Si les erreurs d'échantillonnage sont négligées, on a d'après Eq. (1.151) l'existence d'une unique variable \mathbf{w} dans l'espace de l'ensemble telle que :

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w}. \quad (1.184)$$

De même que Eq. (1.179), la fonction de coût a posteriori fortement contrainte dans l'espace de l'ensemble est :

$$J(\mathbf{w} | \mathbf{y}_{L.}) = \frac{1}{2} \sum_{l=K}^L \|\mathbf{y}_l - \mathcal{F}_l(\mathbf{w})\|_{\mathbf{R}_l^{-1}}^2 + \frac{1}{2} \|\mathbf{w}\|^2, \quad (1.185)$$

avec $\mathcal{F}_l(\mathbf{w}) = \mathcal{H}_l \circ \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w})$ pour $l = K, \dots, L$. À l'itération j de l'analyse variationnelle, les opérateurs $(\mathcal{F}_l)_{l=K, \dots, L}$ sont approchés respectivement par des fonctions affines :

$$\mathcal{F}_l(\mathbf{w}) \simeq \mathcal{F}_l^j(\mathbf{w}), \quad (1.186a)$$

$$\mathcal{F}_l^j(\mathbf{w}) = \mathbf{f}_l^j + \mathbf{F}_l^j(\mathbf{w} - \mathbf{w}^j), \quad (1.186b)$$

pour $l = K, \dots, L$.

De même que dans la section 1.1.8, on donnera à ces approximations affines une portée statistique. Elles engendrent donc une approximation quadratique de la fonction

de coût a posteriori :

$$J(\mathbf{w}|\mathbf{y}_{:L}) \simeq J^j(\mathbf{w}|\mathbf{y}_{:L}), \quad (1.187a)$$

$$J^j(\mathbf{w}|\mathbf{y}_{:L}) = \frac{1}{2} \sum_{l=K}^L \left\| \mathbf{y}_l - \mathcal{F}_l^j(\mathbf{w}) \right\|_{\mathbf{R}_l^{-1}}^2 + \frac{1}{2} \|\mathbf{w}\|^2, \quad (1.187b)$$

qui correspond à une approximation gaussienne de la densité a posteriori :

$$p(\mathbf{w}|\mathbf{y}_{:L}) \simeq p^j(\mathbf{w}|\mathbf{y}_{:L}), \quad (1.188a)$$

$$p^j(\mathbf{w}|\mathbf{y}_{:L}) \propto e^{-J^j(\mathbf{w}|\mathbf{y}_{:L})}. \quad (1.188b)$$

La moyenne \mathbf{w}^{j+1} et la matrice de covariance \mathbf{C}^{j+1} de cette approximation gaussienne se calculent ensuite comme dans la section 1.1.6 avec Eq. (1.63b,1.65a,1.65b) :

$$\mathbf{w}^{j+1} = \mathbf{w}^j - \left(\nabla_{\mathbf{w}}^2 J^j \right)^{-1} \times \nabla_{\mathbf{w}} J^j \left(\mathbf{w}^j | \mathbf{y}_{:L} \right), \quad (1.189a)$$

$$\mathbf{C}^{j+1} = \left(\nabla_{\mathbf{w}}^2 J^j \right)^{-1}. \quad (1.189b)$$

Ce qui constitue une itération de l'analyse variationnelle. Pour évaluer le gradient $\nabla_{\mathbf{w}} J^j(\mathbf{w}^j | \mathbf{y}_{:L})$ et de la hessienne $\nabla_{\mathbf{w}}^2 J^j$, on utilisera la récursivité de la fonction de coût et de ses dérivées :

$$J^j(\mathbf{w}|\mathbf{y}_{l:}) = \frac{1}{2} \left\| \mathbf{y}_l - \mathcal{F}_l^j(\mathbf{w}) \right\|_{\mathbf{R}_l^{-1}}^2 + J^j(\mathbf{w}|\mathbf{y}_{l-1:}) \Rightarrow \quad (1.190a)$$

$$\nabla_{\mathbf{w}} J^j(\mathbf{w}^j | \mathbf{y}_{l:}) = - \left(\mathbf{F}_l^j \right)^{\text{T}} \mathbf{R}_l^{-1} \left(\mathbf{y}_l - \mathbf{f}_l^j \right) + \nabla_{\mathbf{w}} J^j \left(\mathbf{w}^j | \mathbf{y}_{l-1:} \right), \quad (1.190b)$$

$$\nabla_{\mathbf{w}}^2 J^j(\bullet | \mathbf{y}_{l:}) = \left(\mathbf{F}_l^j \right)^{\text{T}} \mathbf{R}_l^{-1} \left(\mathbf{F}_l^j \right) + \nabla_{\mathbf{w}}^2 J^j \left(\bullet | \mathbf{y}_{l-1:} \right), \quad (1.190c)$$

pour $l = K, \dots, L$ avec l'initialisation :

$$\nabla_{\mathbf{w}} J^j \left(\mathbf{w}^j | \mathbf{y}_{K-1:} \right) = \mathbf{w}^j, \quad (1.191a)$$

$$\nabla_{\mathbf{w}}^2 J^j \left(\bullet | \mathbf{y}_{K-1:} \right) = \mathbf{I}_m. \quad (1.191b)$$

Voyons à présent comment linéariser \mathcal{F}_l . Les quantités $\left[\mathbf{f}_l^j \quad \mathbf{F}_l^j \right]$ de Eq. (1.186b) peuvent s'obtenir par résolution d'un problème d'interpolation linéaire :

$$\mathcal{F}_l \left(\mathbf{D}^j \right) = \mathbf{f}_l^j \mathbf{1}_M^{\text{T}} + \mathbf{F}_l^j \left(\mathbf{D}^j - \mathbf{w}^j \mathbf{1}_M^{\text{T}} \right), \quad (1.192)$$

sur un ensemble \mathbf{D}^j de points d'interpolation. Si cet ensemble a pour moyenne empirique

\mathbf{w}^j et pour matrice de déviation \mathbf{W}^j :

$$\mathbf{D}^j = \begin{bmatrix} \mathbf{w}^j & \mathbf{W}^j \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m}\mathbf{U}_M \end{bmatrix}^T, \quad (1.193)$$

la solution du problème d'interpolation est alors :

$$\begin{bmatrix} \mathbf{f}^j & \mathbf{F}^j \end{bmatrix} = \mathcal{F}_l(\mathbf{D}^j) \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} (\mathbf{W}^j)^{-1} \end{bmatrix}. \quad (1.194)$$

Pour économiser des évaluations d'opérateurs, on peut utiliser le schéma suivant pour calculer les $\mathcal{F}_l(\mathbf{D}^j)$:

$$\mathbf{E}_{K-1} = \mathcal{M}_{K-1 \leftarrow 0} (\mathbf{x}_0^b \mathbf{1}_M^T + \mathbf{X}_0^b \mathbf{D}^j), \quad (1.195a)$$

$$\mathbf{E}_l = \mathcal{M}_l(\mathbf{E}_{l-1}), \quad (1.195b)$$

$$\mathcal{F}_l(\mathbf{D}^j) = \mathcal{H}_l(\mathbf{E}_l). \quad (1.195c)$$

Dans les paragraphes suivants, deux définitions de \mathbf{W}^j Eq. (1.193) inspirées de [Sakov et al. \(2012\)](#) et [Bocquet and Sakov \(2014\)](#) sont détaillées.

Méthode « bundle » Si $\mathbf{W}^j = \epsilon \mathbf{I}_m$ alors Eq. (1.194) sont les différences finies de \mathcal{F}_l de pas ϵ :

$$\begin{bmatrix} \mathbf{f}_l^j & \mathbf{F}_l^j \end{bmatrix} \simeq \begin{bmatrix} \mathcal{F}_l(\mathbf{w}^j) & \frac{d\mathcal{F}_l}{d\mathbf{w}}(\mathbf{w}^j) \end{bmatrix}. \quad (1.196)$$

La linéarisation Eq. (1.186b) est alors un développement de Taylor du premier ordre.

Méthode « transform » Si \mathbf{W}^j est telle que $(\mathbf{W}^j)(\mathbf{W}^j)^T = \mathbf{C}^j$, alors Eq. (1.186b) est une estimation de la meilleure approximation affine de \mathcal{F}_l au sens de l'erreur quadratique moyenne selon $p^{j-1}(\mathbf{w}|\mathbf{y}_{:L})$ Eq. (1.188b) :

$$\begin{bmatrix} \mathbf{f}_l^j & \mathbf{F}_l^j \end{bmatrix} \simeq \underset{\begin{bmatrix} \mathbf{f} & \mathbf{F} \end{bmatrix}}{\operatorname{argmin}} \mathbb{E}_{p^{j-1}(\mathbf{w}|\mathbf{y}_{:L})} \left[\left\| \mathcal{F}_l(\mathbf{w}) - \mathbf{f} - \mathbf{F}(\mathbf{w} - \mathbf{w}^j) \right\|^2 \right]. \quad (1.197)$$

En effet, en adaptant les résultats de l'annexe A.3.1, cette solution est :

$$\begin{bmatrix} \mathbb{E}_{p^j(\mathbf{w}|\mathbf{y}_{:L})} [\mathcal{F}_l(\mathbf{w})] & \mathbb{C}_{p^j(\mathbf{w}|\mathbf{y}_{:L})} [\mathcal{F}_l(\mathbf{w}), \mathbf{w}] \mathbb{C}_{p^j(\mathbf{w}|\mathbf{y}_{:L})} [\mathbf{w}]^{-1} \end{bmatrix}, \quad (1.198)$$

et Eq. (1.194) en est une estimation. Ce qui constitue une bonne linéarisation de \mathcal{F}_l dans un sens statistique.

L'analyse variationnelle est arrêtée lorsqu'un critère de convergence du type :

$$\|\mathbf{w}^{j+1} - \mathbf{w}^j\| \ll 0, \quad (1.199)$$

est vérifié. Dans ce cas, on définit $\mathbf{w}^a, \mathbf{W}^a$ tels que :

$$\mathbf{w}^a = \mathbf{w}^*, \quad (1.200)$$

$$(\mathbf{W}^a)(\mathbf{W}^a)^T = \mathbf{C}^*. \quad (1.201)$$

avec l'exposant * se référant au dernier itéré en j . Un ensemble a posteriori déterministe \mathbf{D}^a échantillonnant (approximativement) $p(\mathbf{w}|\mathbf{y}_L)$ est ensuite engendré :

$$\mathbf{D}^a = \begin{bmatrix} \mathbf{w}^a & \mathbf{W}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{M-1}\mathbf{U}_M \end{bmatrix}^T, \quad (1.202)$$

choisir \mathbf{U}_M au hasard permet d'introduire des rotations aléatoires qui améliorent les statistiques de l'ensemble a posteriori (Sakov and Oke, 2008).

Propagation : il s'agit de calculer un ensemble \mathbf{E}_S^b échantillonnant $p(\mathbf{x}_S|\mathbf{y}_L)$ à partir de \mathbf{D}^a . Or :

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w}, \quad (1.203a)$$

$$\mathbf{x}_l = \mathcal{M}_{l \leftarrow l-1}(\mathbf{x}_{l-1}), \quad (1.203b)$$

pour $l = 1, \dots, S$. On engendrera donc \mathbf{E}_S^b par application membre à membre de ces équations :

$$\mathbf{E}_0^b = \mathbf{x}_0^b \mathbf{1}_M^T + \mathbf{X}_0^b \mathbf{D}^a, \quad (1.204a)$$

$$\mathbf{E}_l^b = \mathcal{M}_{l \leftarrow l-1}(\mathbf{E}_{l-1}^b), \quad (1.204b)$$

pour $l = 1, \dots, S$.

Puisque l'IEnKS combine les avantages des méthodes ensemblistes avec ceux des méthodes 4D variationnelles, il est sensé être plus efficace que l'EnKF, l'EnKS ou le 4D-Var. Ceci a été numériquement vérifié sur des modèles de faible dimension (Bocquet and Sakov, 2013, 2014; Haussaire and Bocquet, 2015).

Algorithm 1.1 Un cycle de l'IEEnKS bundle /transform**Require:** Les paramètres de fenêtre $0 \leq K \leq L$, $S = L - K + 1$.Les nouvelles observations $\mathbf{y}_{K:L}$.L'ensemble a priori $\mathbf{E}_0^b \in \mathbb{R}^{n \times M}$.La dimension de l'espace de l'ensemble $m = M - 1$.Une base orthonormale $\begin{bmatrix} \frac{1}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix} \in \mathbb{R}^{M \times M}$.Le pas des différences finies $\epsilon = 1$.Le critère d'arrêt de Gauss-Newton Δ .L'inflation λ .**Ensure:** \mathbf{E}_S^b est l'ensemble a priori du prochain cycle.

1: $\begin{bmatrix} \mathbf{x}_0^b & \mathbf{X}_0^b \end{bmatrix} = \mathbf{E}_0^b \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \end{bmatrix}$

2: $\mathbf{X}_O^b = \lambda \mathbf{X}_0^b$

— Analyse —

3: $\begin{bmatrix} \mathbf{w} & \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_m & \epsilon \mathbf{I}_m \end{bmatrix}$

4: **repeat**

5: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{w} & \mathbf{I}_m \end{bmatrix}$

6: $\mathbf{E}_0 = \begin{bmatrix} \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w} & \mathbf{X}_0^b \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$

7: **for** $l = 0 \dots L$ **do**8: **if** $l > 0$ **then**

9: $\mathbf{E}_l = \mathcal{M}_l(\mathbf{E}_{l-1})$

10: **end if**11: **if** $l \geq K$ **then**

12: $\begin{bmatrix} \mathbf{f} & \mathbf{F} \end{bmatrix} = \mathcal{H}_l(\mathbf{E}_l) \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \mathbf{W}^{-1} \end{bmatrix}$

13: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} += \begin{bmatrix} -\mathbf{F}^T \mathbf{R}_l^{-1} (\mathbf{y}_l - \mathbf{f}) & \mathbf{F}^T \mathbf{R}_l^{-1} \mathbf{F} \end{bmatrix}$

14: **end if**15: **end for**16: solve $\mathbf{C} \times \delta \mathbf{w} = -\mathbf{g}$ for $\delta \mathbf{w} \in \mathbb{R}^m$ 17: solve $\mathbf{W} \mathbf{W}^T = \mathbf{C}^{-1}$ for $\mathbf{W} \in \mathbb{R}^{m \times m}$ 18: $\mathbf{w} += \delta \mathbf{w}$ 19: **until** $\|\delta \mathbf{w}\| \leq \Delta$

— Propagation —

20: $\mathbf{E}_S^b = \mathbf{E}_S$

21: solve $\mathbf{W}^a \mathbf{W}^{aT} = \mathbf{C}^{-1}$ for $\mathbf{W} \in \mathbb{R}^{m \times m}$

22: $\mathbf{E}_0^a = \begin{bmatrix} \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w} & \mathbf{X}_0^b \mathbf{W}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$

23: **for** $l = 1 \dots S$ **do**

24: $\mathbf{E}_l^a = \mathcal{M}_{l \leftarrow l-1}(\mathbf{E}_{l-1}^a)$

25: **end for**

26: $\mathbf{E}_S^b = \mathbf{E}_S^a$

1.3.5 Conclusion de l'assimilation de données ensemble variationnelle fortement contrainte

L'assimilation de données variationnelle d'ensemble fortement contrainte cherche à combiner les méthodes ensemblistes avec les méthodes variationnelles. Les premières permettent d'encoder sur peu de membres les données de la moyenne et de la matrice de covariance d'une densité. Les secondes reposent sur les techniques d'optimisation numériques pour trouver un minimum global de la fonction de coût. La difficulté de produire de tels algorithmes provient des natures différentes de ces méthodes. La première est statistique et porte sur des propriétés globales des densités alors que la seconde est déterministe et porte sur des propriétés locales des densités. Dans l'IEnKS, le passage du local au global se fait en donnant une portée statistique aux approximations affines.

La performance de l'analyse variationnelle est donc liée à la convergence des méthodes d'optimisation itératives. L'IEnKS utilise cette analyse variationnelle pour minimiser une fonction de coût 4D (i.e. s'étalant sur plusieurs pas de temps) dans un cadre séquentiel. Ceci permet différentes stratégies d'estimation de la même densité. Par exemple, l'estimation de $p(\mathbf{x}_2|\mathbf{y}_{:2})$ peut se faire avec un cycle de l'IEnKS $L = 2$ ou bien deux cycles de l'IEnKS $L = 1$. L'étude suivante sur la quasi-stabilité montre que ces stratégies ne sont pas équivalentes en termes de performances lorsque les opérateurs sont non linéaires. Les stratégies avec de plus longues fenêtres sont généralement meilleures si l'on réussit à trouver le minimum global d'une fonction de coût nécessairement moins quadratique.

Chapitre 2

Assimilation de données ensemble variationnelle quasi-statique

2.1 Introduction

2.1.1 Contexte

Lorsque les opérateurs sont non linéaires, la moyenne et la matrice de covariance a posteriori ne peuvent plus être calculées algébriquement. L'alternative est alors d'approcher la densité a posteriori par une densité gaussienne à l'aide d'approximations affines des opérateurs. A son tour, Cette approximation gaussienne permet d'améliorer nos linéarisations. Lorsqu'elles sont basées sur des développements de Taylor, l'itération de ce processus constitue la méthode de Gauss-Newton (GN).

Cependant, la convergence de cette méthode vers un minimum global n'est pas garantie ([Nocedal and Wright, 2006](#)), la méthode pouvant converger vers un extremum local strict ou ne pas converger du tout. Dans l'hypothèse où la masse de probabilité est concentrée dans le voisinage du minimum global, un tel échec conduit à une analyse de mauvaise qualité et, dans un cadre séquentiel, à la divergence de la méthode d'assimilation. Il est donc critique de trouver ce minimum global. Les méthodes d'optimisation spécialisées dans la recherche du minimum global de la fonction de coût sont trop coûteuses pour être une alternative numériquement viable en haute dimension. Il faut alors exploiter la structure de notre problème pour améliorer la convergence vers le minimum global.

2.1.2 Assimilation de données variationnelle quasi-statique

La solution consiste alors à maintenir l'itéré initial de la méthode de GN dans le bassin d'attraction du minimum global. Avec un modèle parfait et chaotique, donc instable et borné, la taille de ce bassin d'attraction semble diminuer avec la longueur de la fenêtre d'assimilation L . Ainsi, le nombre de minimums locaux pouvant leurrer la méthode de GN semble augmenter de manière exponentielle avec L (Pires et al., 1996; Swanson et al., 1998). Ces deux propriétés montrent que la capacité de GN à trouver le minimum global de la fonction de coût diminue avec L . Malheureusement, c'est aussi pour de grandes valeurs de L que les performances d'assimilation sont les meilleures.

Plusieurs stratégies ont été examinées pour dépasser cette restriction. Pires et al. (1996) proposent la minimisation quasi-statique (QS) de la fonction de coût 4D : les observations sont ajoutées graduellement à la fonction de coût, ce qui permet de mettre à jour graduellement l'itéré initial de la méthode variationnelle. Ceci mène à la méthode QSVA pour « quasi-static variational assimilation ». Ye et al. (2015) proposent d'augmenter graduellement les covariances d'erreur modèle dans la fonction de coût 4DVar faiblement contrainte. De cette manière, la non linéarité du modèle est aussi graduellement introduite (Judd et al., 2004). Ye et al. (2015) proposent aussi de paralléliser la minimisation sur plusieurs itérés initiaux pour augmenter les chances de localiser le minimum global.

Le QSVA améliore l'estimation du maximum a posteriori dans le 4DVar. Par contre, le 4DVar traditionnel utilise une matrice de covariance a priori statique, ce qui limite le transfert des statistiques d'un cycle à un autre. Par exemple, dans Pires et al. (1996) il est proposé de décaler graduellement la fenêtre d'assimilation, sans changer de matrice de covariance d'ébauche, pour obtenir une version séquentielle du QSVA.

2.1.3 Méthodes variationnelles d'ensemble

En revanche, les méthodes variationnelles d'ensemble 4D utilisent à la fois une analyse variationnelle non linéaire et une propagation dynamique des statistiques à l'aide des ensembles (Asch et al., 2017). L'amélioration que le QSVA peut apporter à ces méthodes a été suggéré et numériquement évalué par Bocquet and Sakov (2013, 2014). Ce qui motive un traitement analytique et numérique plus approfondi.

Le lisseur de Kalman itératif (IEnKS, Bocquet, 2016; Bocquet and Sakov, 2014) est un archétype de ces méthodes variationnelles d'ensemble 4D où les ensembles sont engendrés de façon déterministe. Il a été montré que l'IEnKS est significativement plus précis en termes de RMSE que le 4DVar et l'EnKS sur des modèles de faibles dimensions.

L'IEEnKS améliore le cyclage de l'assimilation en gérant à la fois la moyenne et la covariance des densités à l'aide d'approximations gaussiennes. Dans le cadre de l'approximation de Laplace, la moyenne a posteriori et la covariance a posteriori sont remplacées par le minimiseur de la fonction de coût et une approximation de l'inverse de la matrice hessienne en ce point respectivement. Ces moments sont ensuite utilisés pour engendrer un ensemble a posteriori de manière déterministe. Cet ensemble a posteriori est ensuite propagé afin d'estimer la moyenne et la matrice de covariance a priori dans le prochain cycle d'assimilation. Il est donc aussi critique pour l'IEEnKS de trouver le minimum global de la fonction de coût.

Goodliff et al. (2015) ont numériquement testé le QSVA sur une collection de méthodes hybrides et EnVar avec le modèle de Lorenz 1963 (Lorenz, 1963) dont ils font varier l'amplitude de la non linéarité. En revanche, l'objet de leur étude n'est pas le cyclage des méthodes d'assimilation et le transfert d'informations entre ces cycles. Ce qui est un aspect central des méthodes EnVar. Au cours de leurs tests, ils ont aussi montré que l'« ensemble transform Kalman smoother » (ETKF) avait les meilleures performances. Méthode que l'IEEnKS est sensée dépasser par construction, ce qui a été numériquement établi dans Bocquet and Sakov (2013). Ce qui renforce le fait que l'IEEnKS puisse être utilisé ici comme archétype des méthodes 4D EnVar déterministes.

2.1.4 Plan

Dans la seconde section de ce chapitre, l'impact des paramètres de fenêtre sur les performances du 4DVar et de l'IEEnKS est examiné. L'accent est mis sur le transfert d'informations d'un cycle à l'autre puisqu'il distingue l'IEEnKS du 4DVar. Pour ce faire, une brève présentation de ces algorithmes est donnée. Une mesure des performances de ces algorithmes est ensuite définie. Elle nous servira à obtenir une expression analytique de la précision de ces méthodes avec un modèle linéaire, diagonal et autonome. Cela nous permet de quantifier l'impact du cyclage et donc des paramètres de fenêtre sur les algorithmes. Après ces préliminaires, le cas non linéaire et chaotique est étudié. Dans la troisième section, le pseudo code de l'IEEnKS quasi-statique (IEEnKS_{QS}) est détaillé. La quatrième section est dédiée aux expériences numériques avec deux modèles de Lorenz de faible dimension et à la réduction du coût numérique de la méthode. Les conclusions sont tirées dans la cinquième section.

Nous soulignons que les développements algorithmiques de cette étude ne sont pas voués à améliorer directement les techniques d'assimilation avec un modèle imparfait ou en haute dimension. Même si Miller et al. (1994) montrent quelques similarités entre les configurations parfaites ou imparfaites d'un modèle de complexité intermédiaire,

l'erreur modèle empêchera, en général, l'utilisation des grandes fenêtres d'assimilation considérées dans cette étude. L'objectif du chapitre est, à la place, de mieux comprendre l'interaction entre la dynamique chaotique, les méthodes d'assimilation variationnelle d'ensemble et leur cyclage.

2.2 La fenêtre et les performances d'assimilation

Après avoir présenté le 4DVar (dans une version avec covariance d'ébauche statique) et l'IEEnKS, nous examinons l'impact des paramètres de fenêtre sur les performances d'assimilation. Cela illustrera l'amélioration introduite par l'IEEnKS par rapport au 4DVar. Nous verrons aussi que, avec un modèle chaotique, plus la fenêtre d'assimilation est longue, meilleure est la performance de ces algorithmes. Ce qui justifie la pertinence du QSVA dans un cadre séquentiel.

Les équations d'observation et d'évolution du système d'assimilation sont supposées de la forme :

$$\mathbf{y}_l = \mathcal{H}(\mathbf{x}_l) + \boldsymbol{\varepsilon}_l, \quad (2.1a)$$

$$\mathbf{x}_{l+1} = \mathcal{M}(\mathbf{x}_l), \quad (2.1b)$$

où l'état inconnu \mathbf{x}_l au temps t_l est propagé au temps t_{l+1} avec la résolvante $\mathcal{M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Le modèle est supposé parfait (pas d'erreurs dans 2.1b) et autonome (\mathcal{M} ne dépend pas du temps). L'opérateur d'observation $\mathcal{H} : \mathbb{R}^n \rightarrow \mathbb{R}^d$ relie l'état \mathbf{x}_l au vecteur d'observations \mathbf{y}_l . Les erreurs d'observation $\boldsymbol{\varepsilon}_l$ sont supposées gaussiennes de moyenne $\mathbf{0}_d \in \mathbb{R}^d$, de matrice de covariance $\mathbf{R} \in \mathbb{R}^{d \times d}$ et décorréelées dans le temps.

2.2.1 Le 4DVar et l'IEEnKS

L'analyse du 4DVar et de l'IEEnKS reposent toutes deux sur une minimisation variationnelle. L'objectif de cette minimisation est de trouver le maximum global de la densité a posteriori $p(\mathbf{x}_0 | \mathbf{y}_{K:L})$ de l'état initial \mathbf{x}_0 à t_0 étant donné les observations $\mathbf{y}_{K:L}$ de t_K à t_L . La densité a posteriori quantifie comment notre connaissance de l'état \mathbf{x}_0 évolue avec les réalisations de $\mathbf{y}_{K:L}$. Ainsi, son maximiseur est le centre du voisinage infinitésimal le plus probable après avoir assimilé ces observations. La fenêtre d'assimilation est illustrée Fig. (1.2). Les paramètres $0 \leq K \leq L$ sont respectivement les indices temporels du début de la fenêtre d'assimilation, du premier et du dernier vecteur d'observation. Le nombre de vecteurs d'observations utilisé dans la fenêtre d'assimilation est donc $L - K + 1$. Pour terminer de spécifier cette densité à posteriori, il faut une hypo-

thèse sur la densité a priori $p(\mathbf{x}_0)$. On la supposera de moyenne $\mathbf{x}_0^b \in \mathbb{R}^n$ et de matrice de covariance $\mathbf{B} \in \mathbb{R}^{n \times n}$:

$$\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0 | \mathbf{x}_0^b, \mathbf{B}). \quad (2.2)$$

A l'aide de ces suppositions, l'expression analytique de la densité a posteriori ou de la fonction de coût associée peut être obtenue. Cette dernière est définie par :

$$G(\mathbf{x}_0 | \mathbf{y}_{K:L}) = -\ln p(\mathbf{x}_0 | \mathbf{y}_{K:L}). \quad (2.3)$$

La notation G est préférée à la notation traditionnelle J pour désigner une fonction de coût exacte, c'est à dire définie comme l'opposée du logarithme d'une densité exacte. L'application de la règle de Bayes donne :

$$G(\mathbf{x}_0 | \mathbf{y}_{K:L}) \propto G(\mathbf{y}_{K:L} | \mathbf{x}_0) + G(\mathbf{x}_0) \quad (2.4)$$

grâce aux hypothèses de gaussianité et d'indépendance sur les erreurs d'ébauche et d'observation on obtient :

$$\begin{aligned} G(\mathbf{x}_0 | \mathbf{y}_{K:L}) &\propto \frac{1}{2} \sum_{l=K}^L \left\| \mathbf{y}_l - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_0) \right\|_{\mathbf{R}^{-1}}^2 \\ &\quad + \frac{1}{2} \left\| \mathbf{x}_0 - \mathbf{x}_0^b \right\|_{\mathbf{B}^{-1}}^2, \end{aligned} \quad (2.5)$$

où l'on rappelle que la notation \propto signifie « égal à une constante additive près » lorsqu'elle est appliquée à des fonctions de coût, $\|\mathbf{x}\|_{\mathbf{S}}^2 = \mathbf{x}^T \mathbf{S} \mathbf{x}$ est le carré de la norme euclidienne associée avec la matrice symétrique, définie-positive \mathbf{S} et \mathcal{M}^l est l compositions du modèle autonome \mathcal{M} .

La propagation correspond à une translation temporelle de S pas de temps. Ainsi, au k ème cycle, la densité a posteriori est $p(\mathbf{x}_{kS} | \mathbf{y}_{K:kS+L})$. Toujours à l'aide du théorème de Bayes, la fonction de coût exacte du k ème cycle est :

$$\begin{aligned} G(\mathbf{x}_{kS} | \mathbf{y}_{K:kS+L}) &\propto \frac{1}{2} \sum_{l=K}^L \left\| \mathbf{y}_{kS+l} - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}) \right\|_{\mathbf{R}^{-1}}^2 \\ &\quad + G(\mathbf{x}_{kS} | \mathbf{y}_{K:(k-1)S+L}) \end{aligned} \quad (2.6)$$

Le terme a priori est $G(\mathbf{x}_{kS} | \mathbf{y}_{K:(k-1)S+L})$. Si le modèle et l'opérateur d'observation sont non linéaires, l'expression analytique de ce terme d'ébauche est difficilement exploitable et doit être simplifiée. Les méthodes du 4DVar et de l'IEEnKS sont basées sur différentes approximations.

Le 4DVar utilise une matrice de covariance d'erreur d'ébauche \mathbf{B} statique, sa fonction de coût au k ème cycle est définie par :

$$J(\mathbf{x}_{kS} | \mathbf{y}_{K:kS+L}) = \frac{1}{2} \sum_{l=K}^L \left\| \mathbf{y}_{kS+l} - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}) \right\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \left\| \mathbf{x}_{kS} - \mathbf{x}_{kS}^b \right\|_{\mathbf{B}^{-1}}^2. \quad (2.7)$$

L'analyse du 4DVar consiste à minimiser cette fonction de coût, on obtient ainsi le minimiseur \mathbf{x}_{kS}^a à t_{kS} . Puisque la fonction de coût dépend des réalisations des observations, \mathbf{x}_{kS}^a est fonction de ces observations, c'est donc une variable aléatoire. Cette analyse est alors propagée au temps $t_{(k+1)S}$ avec la résolvante du modèle pour produire la prochaine ébauche :

$$\mathbf{x}_{(k+1)S}^b = \mathcal{M}^S(\mathbf{x}_{kS}^a). \quad (2.8)$$

En général, $G(\mathbf{x}_{kS} | \mathbf{y}_{K:kS+L})$ et $J(\mathbf{x}_0 | \mathbf{y}_{K:kS+L})$ coïncident seulement au premier cycle d'assimilation grâce à l'hypothèse Eq. (2.2). Ainsi, le terme a priori $\frac{1}{2} \left\| \mathbf{x}_{kS} - \mathbf{x}_{kS}^b \right\|_{\mathbf{B}^{-1}}^2$ n'est qu'une approximation quadratique du terme a priori exact $G(\mathbf{x}_{kS} | \mathbf{y}_{K:(k-1)S+L})$. Par définition, la matrice de covariance d'erreur a priori \mathbf{B} est la même à chaque cycle d'assimilation, ce qui n'est pas le cas de l'IEnKS.

L'IEnKS (Bocquet and Sakov, 2014) est une méthode d'ensemble avec une analyse variationnelle. Deux versions de l'algorithme existent : la version « singular data assimilation » (SDA) où les mêmes observations ne sont assimilées qu'une seule fois, et la version « multiple data assimilation » (MDA) où les mêmes observations sont assimilées plusieurs fois. Nous nous concentrerons sur la version SDA lors de ces développements théoriques. La version MDA, qui peut être vue comme le premier schéma EnVar quasi-statique, sera utilisée dans la partie numérique comme référence. Notons que dans l'IEnKS SDA le nombre d'observations par fenêtre vérifie :

$$L - K + 1 = S. \quad (2.9)$$

Au k ème cycle de l'IEnKS, l'ensemble a priori au temps t_{kS} est obtenu par propagation depuis le cycle précédent. Les membres de l'ensemble sont les colonnes de la matrice \mathbf{E}_{kS}^b qui est vue comme une matrice aléatoire à valeurs dans $\mathbb{R}^{n \times M}$. Elle est utilisée pour

estimer la moyenne et la matrice de covariance a priori :

$$\mathbb{E}_p(\mathbf{x}_{kS} | \mathbf{y}_{K:(k-1)S+L}) [\mathbf{x}_{kS}] \simeq \mathbf{x}_{kS}^b, \quad (2.10a)$$

$$\mathbb{C}_p(\mathbf{x}_{kS} | \mathbf{y}_{K:(k-1)S+L}) [\mathbf{x}_{kS}] \simeq (\mathbf{X}_{kS}^b) (\mathbf{X}_{kS}^b)^\top, \quad (2.10b)$$

où \mathbf{x}_{kS}^b et \mathbf{X}_{kS}^b sont la moyenne et une matrice de déviation de \mathbf{E}_{kS}^b respectivement :

$$\begin{bmatrix} \mathbf{x}_{kS}^b & \mathbf{X}_{kS}^b \end{bmatrix} = \mathbf{E}_{kS}^b \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{M-1}} \mathbf{U}_M \end{bmatrix}, \quad (2.11)$$

avec $\begin{bmatrix} \frac{1}{\sqrt{M}} & \mathbf{U}_M \end{bmatrix}$ une base orthonormale de \mathbb{R}^M de premier vecteur $\begin{bmatrix} \frac{1}{\sqrt{M}} & \dots & \frac{1}{\sqrt{M}} \end{bmatrix}^\top$. Les équations (2.10a,2.10b) sont des approximations à cause des erreurs d'échantillonnage. Si le vecteur d'état est de la forme :

$$\mathbf{x}_{kS} = \mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w}_{kS}, \quad (2.12)$$

avec $\mathbf{w}_{kS} \in \mathbb{R}^m$, $m = M - 1$ la variable d'ensemble, la fonction de coût a posteriori de cette variable au k ème cycle de l'IEnKS est :

$$J(\mathbf{w}_{kS} | \mathbf{y}_{K:kS+L}) = \frac{1}{2} \sum_{l=K}^L \left\| \mathbf{y}_{kS+l} - \mathcal{F}_l(\mathbf{w}_{kS}) \right\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathbf{w}_{kS}\|^2, \quad (2.13a)$$

$$\mathcal{F}_l(\mathbf{w}_{kS}) = \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w}_{kS}), \quad (2.13b)$$

où $\|\bullet\|^2 = \|\bullet\|_{\mathbf{I}_m}^2$ avec \mathbf{I}_m la matrice identité de \mathbb{R}^m . L'analyse de l'IEnKS consiste à minimiser cette fonction de coût, on obtient ainsi un ensemble a posteriori \mathbf{E}_{kS}^a au temps t_{kS} ayant pour moyenne et matrice de déviation :

$$\mathbf{x}_{kS}^a = \mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w}_{kS}^a, \quad (2.14a)$$

$$\mathbf{X}_{kS}^a = \mathbf{X}_{kS}^b \mathbf{W}_{kS}^a, \quad (2.14b)$$

respectivement, avec \mathbf{w}_{kS}^a le minimiseur de la fonction de coût et \mathbf{W}_{kS}^a une racine de l'inverse de la matrice hessienne en ce minimum approchée dans la méthode de Gauss-

Newton par :

$$\nabla^2 J(\mathbf{w}_{kS}^a | \mathbf{y}_{K:kS+L}) \simeq \mathbf{I}_m + \sum_{l=K}^L \mathbf{F}_l^T \mathbf{R}^{-1} \mathbf{F}_l, \quad (2.15a)$$

$$\mathbf{F}_l = \frac{d\mathcal{F}_l}{d\mathbf{w}_{kS}^a}(\mathbf{w}_{kS}^a), \quad (2.15b)$$

pour éviter de calculer les dérivées secondes du modèle. Cet ensemble a posteriori \mathbf{E}_{kS}^a est ensuite propagé au temps $t_{(k+1)S}$ pour produire l'ensemble a priori du cycle suivant :

$$\mathbf{E}_{(k+1)S}^b = \mathcal{M}^S(\mathbf{E}_{kS}^a). \quad (2.16)$$

Le cycle d'assimilation est ainsi complété. Les méthodes du 4DVar et de l'IEEnKS sont schématisées dans la figure (2.1)

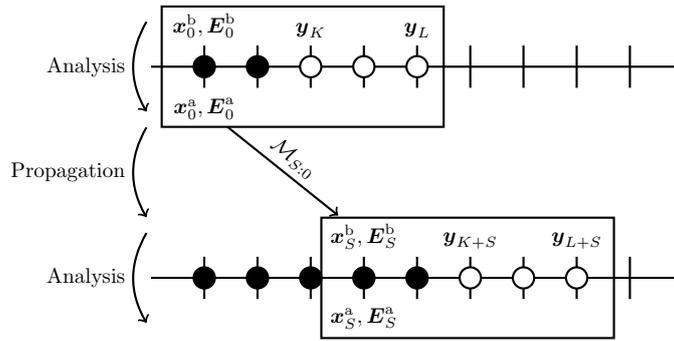


FIGURE 2.1 – Enchaînement des deux premiers cycles du 4DVar et de l'IEEnKS avec $S = 3, L = 4$. La première analyse du 4DVar utilise l'ébauche \mathbf{x}_0^b à t_0 et les observations $\mathbf{y}_{L:K}$ pour produire l'analyse \mathbf{x}_0^a à t_0 . Elle est ensuite propagée de S pas de temps pour produire la nouvelle ébauche à t_S . L'IEEnKS fait de même avec un ensemble. Le rectangle symbolise la fenêtre d'assimilation courante, les points blancs représentent les nouvelles observations assimilées pendant le cycle courant, les points noirs représentent les observations déjà assimilées, les barres sans points représentent les observations qui ne sont pas encore assimilées.

2.2.2 Performance de l'assimilation

On s'intéresse ici à deux mesures permettant d'évaluer l'efficacité du 4DVar et de l'IEEnKS en fonction des paramètres S et L : la RMSE empirique et un équivalent théorique.

Pendant le k ème cycle, l'algorithme engendre une analyse \mathbf{x}_{kS}^a au temps t_{kS} à l'aide

des observations. Cette analyse est ensuite propagée de l pas de temps pour obtenir l'analyse au temps \mathbf{x}_{kS+l}^a sensée approcher l'état du système \mathbf{x}_{kS+l} . La « root mean square error » (RMSE) est la mesure traditionnelle de la performance d'assimilation. Elle est définie par :

$$\text{RMSE} = \frac{1}{\sqrt{n}} \left\| \mathbf{x}_{kS+l} - \mathbf{x}_{kS+l}^a \right\|.$$

La RMSE prend des noms différents en fonction du moment où elle est calculée. Si $l = L$, il s'agit de la RMSE de filtrage ; si $l = L$, il s'agit de la RMSE de lissage ; dans les autres cas, il s'agit de la RMSE avec lag $L - l$.

La RMSE dépend des réalisations de variables aléatoires, c'est donc aussi une variable aléatoire. Dans nos expériences numériques, la RMSE est moyennée sur les cycles d'assimilation pour atténuer cette variabilité :

$$\text{aRMSE}_N = \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{\sqrt{m}} \left\| \mathbf{x}_{kS+l} - \mathbf{x}_{kS+l}^a \right\|. \quad (2.17)$$

Si on suppose l'existence d'un couple $(\mathbf{x}_{\infty S+l}, \mathbf{x}_{\infty S+l}^a)$ dont la densité est invariante et ergodique par rapport au décalage temporel :

$$T : (\mathbf{x}_{kS+l}, \mathbf{x}_{kS+l}^a) \mapsto (\mathbf{x}_{(k+1)S+l}, \mathbf{x}_{(k+1)S+l}^a), \quad (2.18)$$

alors, d'après le théorème ergodique de Birkhoff (Walters, 1982), la série $(\text{aRMSE}_N)_N$ converge quand $N \rightarrow \infty$ et sa limite aRMSE vérifie :

$$\text{aRMSE} = \frac{1}{\sqrt{m}} \mathbb{E}_{p(\mathbf{x}_{\infty S+l}, \mathbf{x}_{\infty S+l}^a)} \left[\left\| \mathbf{x}_{\infty S+l} - \mathbf{x}_{\infty S+l}^a \right\|^2 \right]. \quad (2.19)$$

Dans ce cas, la aRMSE mesure l'impact à long terme du cyclage sur la performance d'assimilation. Cependant, cette limite est difficile à manipuler algébriquement. C'est pourquoi l'espérance de la MSE (eMSE) sera préférée dans les développements théoriques. Elle sera notée :

$$P_{kS+l} = \mathbb{E}_{p(\mathbf{x}_{kS+l}, \mathbf{x}_{kS+l}^a)} \left[\left\| \mathbf{x}_{kS+l} - \mathbf{x}_{kS+l}^a \right\|^2 \right]. \quad (2.20)$$

Dans la section suivante, nous nous intéresserons à l'effet à long terme du cyclage sur P_{kS+l} . Des hypothèses simplificatrices seront faites pour obtenir une expression analytique de :

$$P_{\infty S+l} = \lim_{k \rightarrow \infty} P_{kS+l}, \quad (2.21)$$

en tant que fonction de S et L .

2.2.3 Performance dans le cas linéaire, diagonal et autonome

Pour obtenir une expression analytique de la eMSE du 4DVar et de l'IEEnKS, des hypothèses hautement simplificatrices vont être faites.

D'abord, le modèle est supposé être la résolvante d'une équation différentielle linéaire, diagonale, autonome et ordinaire. Il vérifie alors :

$$\mathcal{M}^l(\mathbf{x} + \delta\mathbf{x}) = \mathcal{M}^l(\mathbf{x}) + \mathbf{M}^l\delta\mathbf{x}, \quad (2.22)$$

avec $\mathbf{M} = \text{diag}(\alpha_1, \dots, \alpha_n)$ la partie linéaire de \mathcal{M} supposée diagonale et indépendante de \mathbf{x} . Nous supposons de plus $\mathcal{H} = h\mathbf{I}_n$, $\mathbf{B} = b\mathbf{I}_n$, $\mathbf{R} = r\mathbf{I}_n$ avec \mathbf{I}_n la matrice identité de \mathbb{R}^n et $h, r, b > 0$. Avec ces hypothèses, l'annexe B.1 fournit une expression de la eMSE asymptotique dans le cas uni-varié. La généralisation au cas diagonal multi-varié est obtenue en additionnant les eMSEs de chaque direction :

$$P_{\infty S+l}^{\text{4D-Var}} = \sum_{i=1}^n \begin{cases} \infty & \text{si } \Delta_i \geq 1 \\ \frac{b^2 \Sigma_{K,i}^L}{\alpha_i^{2(S-1)} (1-\Delta_i)} & \text{sinon} \end{cases}, \quad (2.23a)$$

$$\Sigma_{K,i}^L = \frac{h^2}{r} \frac{\alpha_i^{2(L+1)} - \alpha_i^{2K}}{\alpha_i^2 - 1}, \quad (2.23b)$$

$$\Delta_i = \frac{\alpha_i^{2S}}{(1 + b \Sigma_{K,i}^L)^2}. \quad (2.23c)$$

Le cas $\Delta_i \geq 1$ signifie que trop de crédit est donné à la variance a priori qui est approchée par la constante b dans le 4DVar présenté ici. Ainsi, l'information portée par les observations n'est pas suffisante pour atténuer la croissance exponentielle des erreurs lors de la propagation.

En ce qui concerne l'IEEnKS, les matrices de déviation sont supposées être de rang plein pour ne pas avoir de matrices de covariance singulières. De plus, la linéarité du modèle est utilisée pour propager les statistiques d'ébauche :

$$\mathbf{x}_{(k+1)S}^b = \mathcal{M}^S(\mathbf{x}_{kS}^a), \quad (2.24a)$$

$$\mathbf{X}_{(k+1)S}^b = \mathbf{M}^S \mathbf{X}_{kS}^a, \quad (2.24b)$$

$$\mathbf{X}_0^b = \mathbf{B}^{\frac{1}{2}}. \quad (2.24c)$$

Les erreurs d'échantillonnage sont ainsi évitées puisqu'elles ne sont pas le sujet de cette

étude. Cet IEnKS simplifié est en fait un lisseur de Kalman (Bocquet and Carrassi, 2017; Cosme et al., 2012). Avec ces suppositions, l'annexe B.2 fournit une expression de la eMSE asymptotique de l'IEnKS dans le cas uni-varié. L'optimalité de la eMSE de l'IEnKS est aussi démontrée. La généralisation au cas diagonal multi-varié est aussi obtenue en ajoutant les eMSEs dans chaque direction :

$$P_{\infty S+l}^{\text{IEnKS}} = \sum_{i=1}^n \begin{cases} 0 & \text{si } |\alpha_i| \leq 1 \\ \frac{r}{h^2 \alpha_i^{2(L-l)}} \frac{\alpha_i^2 - 1}{\alpha_i^2} & \text{sinon} \end{cases}. \quad (2.25)$$

Cette expression montre que les composantes de la eMSE dans les directions stables sont nulles. En effet, l'IEnKS est au moins plus performant qu'une intégration du modèle sans assimilation dont l'erreur dans les directions stables tend vers zéro¹. Ce n'est pas le cas pour la eMSE du 4DVar Eq. (2.23a), puisque la covariance a priori statique réintroduit continuellement une fausse variance dans les directions stables. Dans Trevisan et al. (2010), les variances d'erreur dans les directions stables sont mises à zéro pour augmenter la précision de l'assimilation.

Dans ce qui suit, la dépendance de la eMSE avec les paramètres de fenêtre S et L pour l'IEnKS et le 4DVar est étudiée. Nous nous focaliserons sur un cas bi-varié avec $\alpha_1 = 1.2$ et $\alpha_2 = 0.8$ pour avoir une direction stable et une instable avec $h = b = r = 1$. A l'aide des Eq. (2.23a, 2.25), les eMSEs asymptotiques de lissage et de filtrage sont tracés en fonction de L et S dans la figure (2.2) avec leurs composantes dans la direction stable et instable. Ces graphes sont interprétés dans ce qui suit.

L'expression de la eMSE asymptotique de l'IEnKS est de la forme :

$$P_{\infty S+l}^{\text{IEnKS}} = c_1 \alpha_1^{2(l-L)}, \quad (2.26)$$

où c_1 ne dépend pas de S, L, l . La eMSE asymptotique de l'IEnKS ne dépend donc que du lag $L - l$ et ne dépend pas de S , la contribution de la direction stable est nulle. La eMSE de filtrage ($l = L$) est donc constante : la propagation compense l'analyse. La eMSE de lissage ($l = 0$) décroît exponentiellement avec L : l'assimilation d'observations futures améliore les performances.

En ce qui concerne le 4DVar, on supposera K fixé et $S \rightarrow \infty$ pour simplifier l'expression de la eMSE :

$$P_{\infty S+l}^{\text{4D-Var}} = P_{\infty S+l}^{\text{IEnKS}} (1 + o(1)) + c_2 \alpha_2^{2l} (1 + o(1)), \quad (2.27)$$

1. Le fait que l'erreur d'analyse se trouve dans le sous-espace instable du modèle est plus générale (Bocquet and Carrassi, 2017).

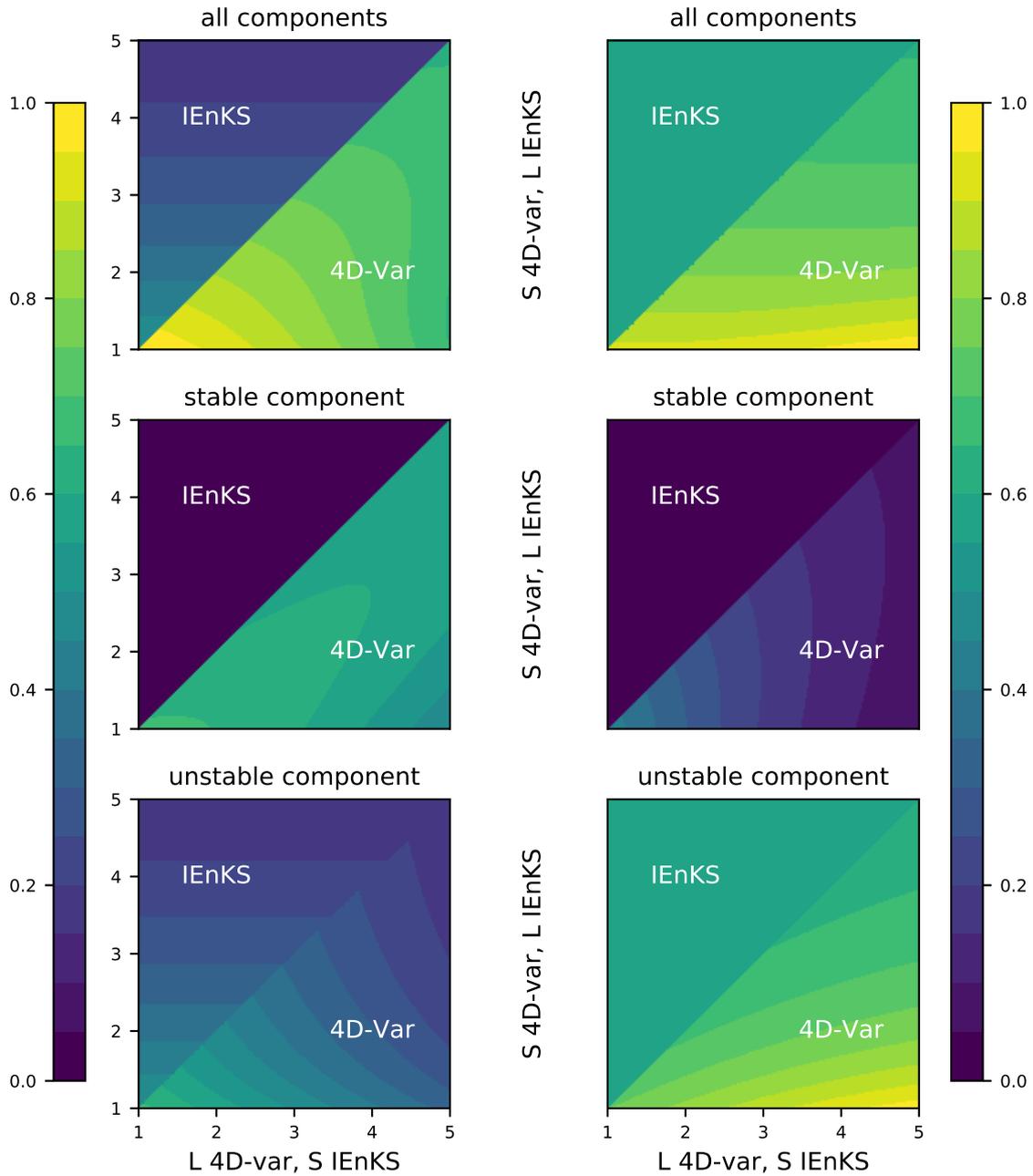


FIGURE 2.2 – Les eMSEs asymptotiques de lissage (a) et de filtrage (b) du 4DVar et de l'IEnKS tracés en fonction de L et S avec leurs composantes dans la direction stable et instable. Le paramètre L est sur l'axe des abscisses pour le 4DVar et sur l'axe des ordonnées pour l'IEnKS. Le paramètre S est sur l'axe des abscisses pour l'IEnKS et sur l'axe des ordonnées pour le 4DVar.

où c_2 est une constante de S, l et $o(1) \rightarrow 0$ quand $S \rightarrow \infty$. La composante instable est proche de la eMSE de l'IEnKS. La plus grande différence concerne la composante stable. La modélisation inexacte de la covariance a priori dans le 4DVar ajoute un terme pénalisant à sa eMSE.

Pour qualifier l'effet à long terme du cyclage sur la performance, la eMSE de filtrage est plus instructive. En effet, la eMSE de lissage s'améliore avec L puisqu'elle tient compte d'observations ultérieures au temps de l'analyse. Cette amélioration domine l'effet potentiellement nuisible de l'approximation gaussienne de la densité a priori. Dans la eMSE de filtrage, cette amélioration est balancée par la propagation de l'analyse à la fin de la fenêtre. Dans la figure (2.2), la composante stable de la eMSE de filtrage est atténuée et a donc peu d'effet. Cependant, le paramètre S améliore la composante instable de la eMSE. Plus S est grand, plus $P_{\infty S+l}^{4D-Var}$ est proche de $P_{\infty S+l}^{IEnKS}$ qui est optimal (Cf annexe B.2). Une explication qualitative est que pour assimiler le même nombre d'observations, un 4DVar avec une plus grande valeur de S demande moins de cycles et invoque donc moins fréquemment l'approximation gaussienne de la densité a priori, ce qui rend l'analyse plus correcte.

La figure (2.3) trace les eMSEs asymptotiques des deux algorithmes en fonction du lag $L - l$ pour $S = L = 5$. Ces courbes sont similaires à celles de Trevisan et al. (2010). La eMSE du 4DVar peut être mise sous la forme d'une fonction du lag $L - l$:

$$P_{\infty S+l}^{4D-Var} \simeq c_1 \alpha_1^{-2\text{lag}} + c_2 \alpha_2^{2(L-\text{lag})}, \quad (2.28)$$

ainsi, la composante instable de la eMSE est une fonction exponentiellement décroissante du lag et la composante stable une fonction exponentiellement croissante du lag. La somme est donc décroissante lorsque la composante instable domine et croissante lorsque la composante stable domine.

Dans cette section, nous avons étudié la précision de l'IEnKS et du 4DVar cyclés en tant que fonction des paramètres de fenêtre dans un contexte linéaire, autonome et diagonal. Nous avons montré que le paramètre L améliore les performances de lissage et que le paramètre S améliore les performances de filtrage. Ces propriétés seront étudiées numériquement dans un contexte non linéaire dans la prochaine section.

2.2.4 Performance dans le cas non linéaire et chaotique

Les résultats de la section 2.2.3 encouragent à choisir L grand pour améliorer les eMSEs de lissage du 4DVar et de l'IEnKS. Pour le filtrage, la propagation des erreurs à la fin de la fenêtre d'assimilation balance le gain de eMSE dû à l'assimilation d'observations

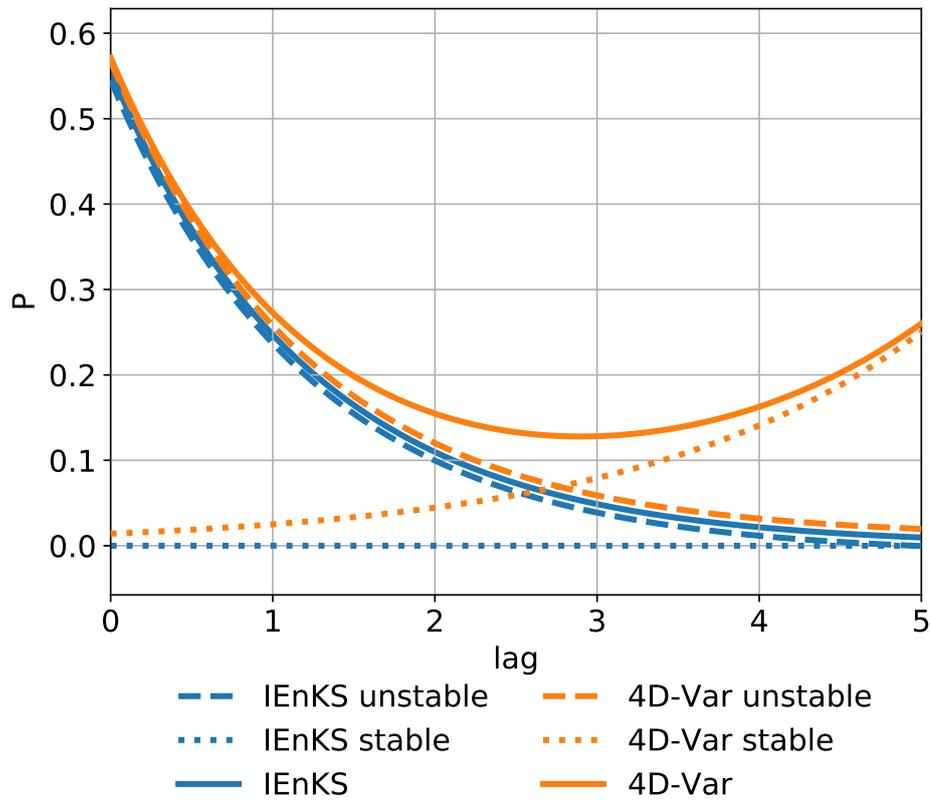


FIGURE 2.3 – Les eMSEs asymptotiques du 4DVar et de l'IEnKS en fonction du lag $L - l$. Les courbes se superposant ont été légèrement translatées pour être lisibles et $S = L = 5, h = b = r = 1$.

futures. Ainsi, la eMSE de filtrage n'est pas améliorée par l'assimilation d'observations lointaines dans le temps, elle est plutôt affectée par les approximations de la densité a priori. Pour assimiler le même nombre d'observations, les méthodes utilisant un plus grand paramètre S nécessitent moins de cycles. Par conséquent, elles reposent moins souvent sur ces approximations de la densité a priori. C'est pourquoi les performances de filtrage du 4DVar s'améliorent avec S . Pour l'IEEnKS, comme dans la section 2.2.3, l'approximation de la densité a priori était exacte, il n'y avait donc pas de dépendance des performances avec S . Ce n'est plus vrai dans un contexte non-linéaire où l'utilisation de densités gaussiennes a priori devient une approximation. Donc, par analogie avec le 4DVar dans la section précédente, on s'attend à une amélioration des performances de filtrage de l'IEEnKS avec S .

2.2.4.1 Des minimums locaux multiples

Avec un modèle chaotique, Pires et al. (1996) montrent que le nombre d'extremum locaux de la fonction de coût du 4DVar augmente avec L , rendant la minimisation de plus en plus problématique.

Ce comportement sera illustré par le modèle de Lorenz 95 (L95 Lorenz and Emanuel, 1998). Il représente un cercle d'atmosphère de latitude moyenne et est décrit par un ensemble de n équations différentielles non linéaires :

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \quad (2.29)$$

où x_i est la i ème composante modulo n de \mathbf{x} , $n = 40$ et $F = 8$. Cette équation est intégrée à l'aide d'un schéma de Runge-Kutta d'ordre quatre avec un pas de temps $\delta t = 0.05$. La dynamique de L95 est chaotique et son plus grand exposant de Lyapunov est $\lambda \simeq 1.7$.

La figure 2.4 montre des profils typiques de la fonction de coût de l'IEEnKS dans une direction de l'ensemble analysé pour différentes valeurs de S, L . Le système est observé à chaque pas de temps et $\mathcal{H} = \mathbf{B} = \mathbf{R} = \mathbf{I}_n$. Les courbes ont de plus en plus d'extremum locaux lorsque L augmente et leurs replis ont de moins en moins d'amplitude lorsque S augmente (un effet de moyenne semble apparaître avec le nombre d'observations).

Cet aspect accidenté cause des problèmes de minimisation. Une telle procédure de minimisation pour l'IEEnKS est la méthode de Gauss-Newton (GN, e.g. Björck, 1996). Ce n'est pas une procédure globale dans le sens où, en fonction du point de départ, la méthode peut converger vers un extremum local, demander de nombreuses itérations ou même diverger. Cependant, si la fonction de coût est quadratique, le minimum global

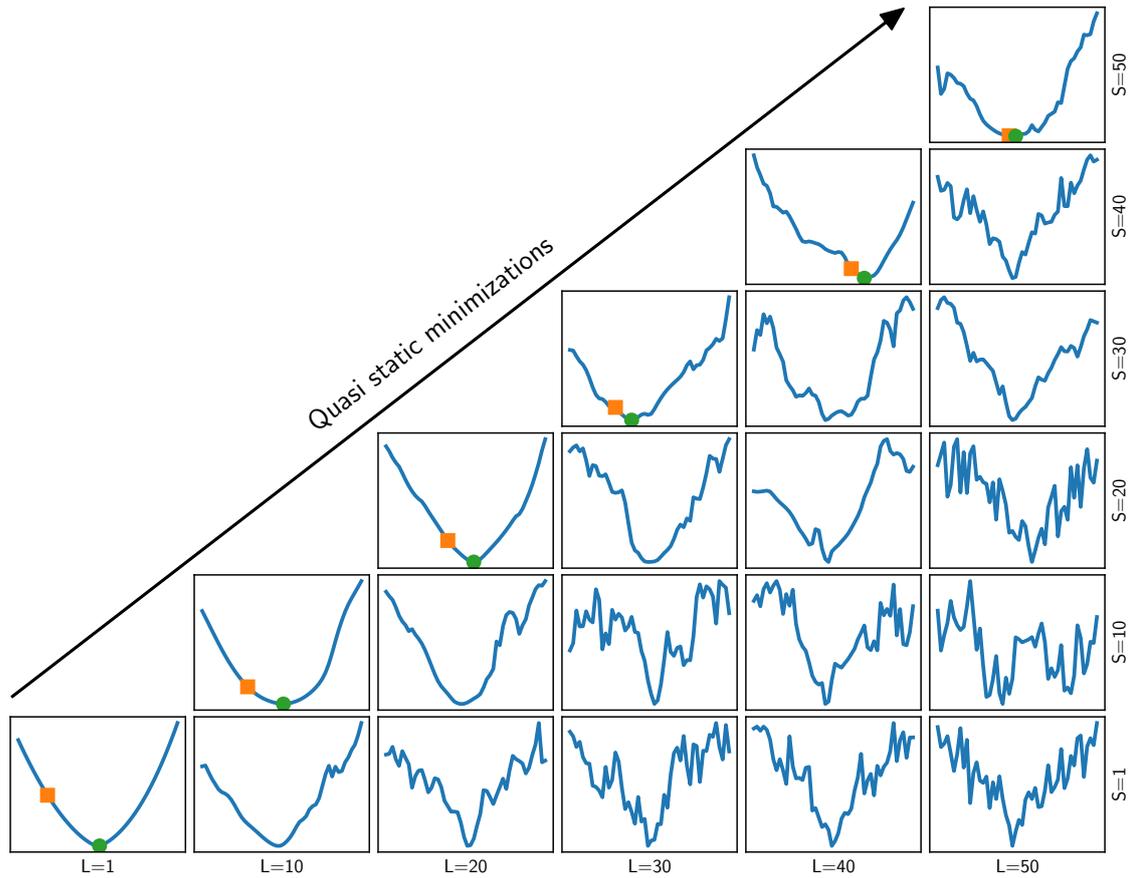


FIGURE 2.4 – Fonctions de coût de l’IEnKS projetées dans une direction de l’ensemble analysé (donc centrées et normalisées) avec différents paramètres S et L . Une minimisation quasi-statique est schématisée sur ces panneaux. Depuis la fonction de coût en bas à gauche (le carré orange est le point de départ de la minimisation, le point vert est le minimum) jusqu’à la fonction de coût en haut à droite, des groupes de 9 puis 10 vecteurs d’observations sont progressivement ajoutés à la fenêtre d’assimilation, le point de départ de la minimisation suivante étant défini par le minimum de la minimisation courante.

est atteint en une seule itération. La non quadraticité de la fonction de coût est induite par la non linéarité du modèle. Dans la suite, nous donnerons un argument heuristique permettant d'obtenir une borne sur le paramètre S au-delà de laquelle la méthode de Gauss-Newton manque probablement le minimum global. Dans une certaine mesure, cet argument peut être vu comme une amélioration de la notion de fenêtre d'assimilation utile (Useful DAW) de Pires et al. (1996), c'est à dire la fenêtre au-delà de laquelle le gain de performance est négligeable. A la différence de la fenêtre d'assimilation utile, nous tenons compte dans la suite du cyclage et de la non linéarité.

2.2.4.2 Longueur effective de la fenêtre d'assimilation

D'abord, les propriétés de convergence de la méthode de Gauss-Newton sont simplifiées. On supposera que la méthode converge vers le minimiseur global si et seulement si l'itéré initial se trouve dans un voisinage du minimiseur où la fonction de coût de l'IEnKS est pratiquement quadratique.

Malheureusement, ce minimiseur est inconnu puisque la fonction de coût dépend de la réalisation de nombreuses variables aléatoires. Pires et al. (1996) introduisent pour cela une fonction de coût sans erreur (error-free cost function). Nous choisirons ici une fonction de coût en moyenne $J_{\infty S}$ définie par :

$$J_{\infty S}(\mathbf{w}) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} J(\mathbf{w} | \mathbf{y}_{kS+L:K}). \quad (2.30)$$

En s'appuyant sur des hypothèses d'ergodicité, l'annexe B.3 prouve que cette fonction de coût moyenne vérifie :

$$J_{\infty S}(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{dS}{2} + \frac{1}{2} \mathbb{E} \left[\|\delta \mathbf{x}_{\infty S+l}^b\|^2 \right], \quad (2.31a)$$

$$\delta \mathbf{x}_{\infty S+l}^b = \mathcal{M}^l(\mathbf{x}_{\infty S}^b + \mathbf{X}_{\infty S}^b \mathbf{w}) - \mathcal{M}^l(\mathbf{x}_{\infty S}), \quad (2.31b)$$

où les variables ergodiques $\mathbf{x}_{\infty S}^b$, $\mathbf{X}_{\infty S}^b$ et $\mathbf{x}_{\infty S}$ ont été définies dans la section 2.2.2 et l'annexe B.3, l'espérance est prise suivant ces variables. D'après Eq. (2.31a,2.31b), une condition suffisante pour que l'itéré initial $\mathbf{w} = \mathbf{0}$ se trouve dans un voisinage du minimiseur global où la fonction de coût est pratiquement quadratique est que $\mathbf{x}_{\infty S}^b$ soit dans un voisinage de $\mathbf{x}_{\infty S}$ où tous les $(\mathcal{M}^l)_{K \leq l \leq L}$ sont pratiquement linéaires.

Si le comportement du modèle est pratiquement linéaire instable et diagonal, on peut utiliser Eq. (2.25) pour estimer les termes de la somme Eq. (2.31a) au point initial

$\mathbf{w} = \mathbf{0}$:

$$\mathbb{E} \left[\left\| \delta \mathbf{x}_{\infty S+l}^b \right\|^2 \right] \simeq n \alpha^{2(S+l-L)} \frac{\alpha^2 - 1}{\alpha^2}, \quad (2.32)$$

où α est un élément diagonal du modèle que l'on a supposé être le même dans chaque direction de l'espace et le S du membre de droite tient compte de la propagation. Cependant dans un système physique borné, le membre de droite d'Eq. (2.32) ne peut pas augmenter indéfiniment avec $l + S$. Une telle saturation du modèle impose :

$$\mathbb{E} \left[\left\| \delta \mathbf{x}_{\infty S+l}^b \right\|^2 \right] \leq B, \quad (2.33)$$

avec B une certaine borne. Ainsi, les équations (2.32,2.33) permettent d'obtenir l'inégalité suivante sur S :

$$S \leq S_{\max}, \quad (2.34a)$$

$$S_{\max} = \frac{\ln(B) - \ln(1 - \alpha^{-2})}{2 \ln(\alpha)}. \quad (2.34b)$$

Nous avons fait le choix $l = L$ puisqu'il correspond au cas le plus contraignant. Quand l'inégalité 2.34a est violée, $\mathbf{x}_{\infty S}^b$ est suffisamment loin de $\mathbf{x}_{\infty S}$ pour que la non linéarité de \mathcal{M}^L soit significative.

Pour pouvoir appliquer cette inégalité au modèle de L95, nous faisons le choix suivant :

$$\alpha = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \sigma \left(\frac{d\mathcal{M}}{d\mathbf{x}}(\mathbf{x}_k) \right),$$

où $\sigma \left(\frac{d\mathcal{M}}{d\mathbf{x}}(\mathbf{x}_k) \right)$ est la moyenne des valeurs singulières plus grande que 1 de $\frac{d\mathcal{M}}{d\mathbf{x}}(\mathbf{x}_k)$. Cette quantité correspond à l'amplification moyenne des erreurs par $\frac{d\mathcal{M}}{d\mathbf{x}}(\mathbf{x}_k)$ sur le sous espace-instable local. Nous choisissons aussi la moyenne du carré de la norme entre 2 longues trajectoires du modèle pour B :

$$B = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left\| \mathbf{x}_{kS} - \mathcal{M}_{kS}(\mathbf{x}_0^b) \right\|^2. \quad (2.35)$$

Cette quantité est plus grande que $\mathbb{E} \left[\left\| \delta \mathbf{x}_{\infty S+l}^b \right\|^2 \right]$ parce que les performances asymptotiques de l'IEnKS sont au moins meilleures qu'une intégration libre du modèle sans assimilation. Avec ces valeurs de α et B , on trouve $S_{\max} = 14$.

La figure 2.5 montre les aRMSEs de filtrage et lissage d'un IEnKS $L = S$ avec L95 en fonction de S . Les performances se détériorent fortement pour $S > 16$ ce qui est cohérent avec notre estimation. Cette figure montre aussi une autre différence avec le cas linéaire :

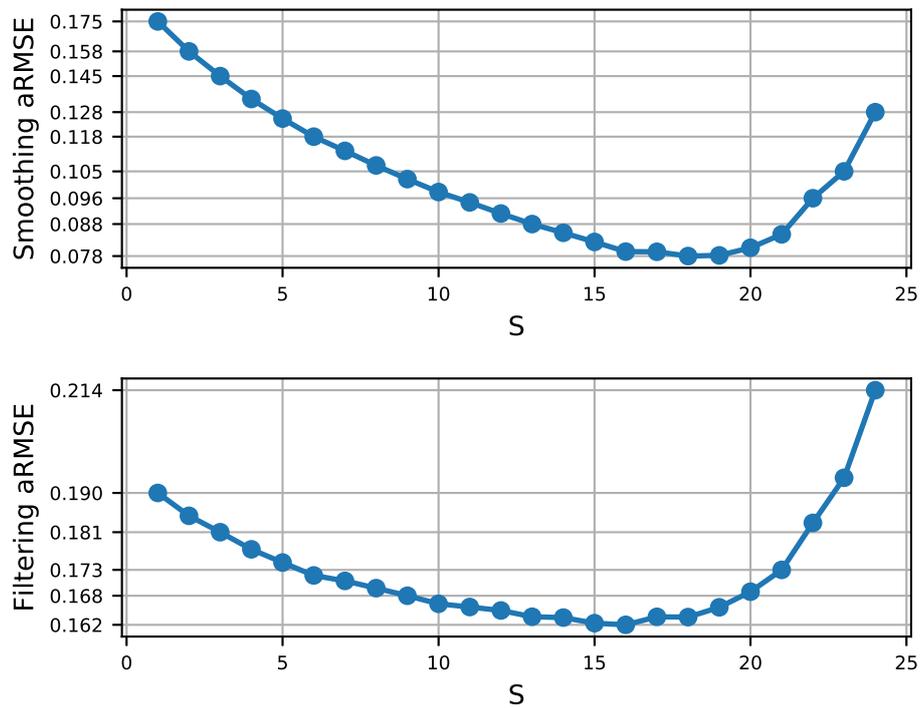


FIGURE 2.5 – aRMSE de lissage (a) et de filtrage (b) en échelle logarithmique d'un IEnKS (5×10^5 cycles) avec Gauss-Newton $L = S$ en fonction de S avec le modèle de L95. La version « finite-size » [Bocquet \(2011\)](#) a été utilisée pour tenir compte des erreurs d'échantillonnage et éviter de régler l'inflation. $\mathcal{H} = \mathbf{R} = \mathbf{B} = \mathbf{I}_n$, $m = 20 - 1$.

la aRMSE de filtrage dépend de S . La discussions précédente sur la convergence de GN explique cette dépendance pour de grandes valeurs de S . Cependant, pour de petites valeurs, la décroissance de la aRMSE n'est pas encore expliquée. C'est une conséquence de l'approximation gaussienne de la densité a priori. À chaque cycle, l'IEnKS utilise l'ensemble \mathbf{E}_{kS}^b pour estimer les deux premiers moments de cette densité a priori et fait l'approximation suivante :

$$p(\mathbf{x}_{kS} | \mathbf{y}_{(k-1)S+L:K}) \simeq \mathcal{N}\left(\mathbf{x}_{kS} | \mathbf{x}_{kS}^b, (\mathbf{X}_{kS}^b)(\mathbf{X}_{kS}^b)^T\right). \quad (2.36)$$

La conséquence d'une telle approximation est une perte d'information. Plus elle est faite, plus la fonction de coût de l'IEnKS s'éloigne de G . C'est exactement ce qu'il se passe lorsque S est petit : pour assimiler le même nombre d'observations, l'IEnKS utilise plus de cycles, il fait donc l'approximation gaussienne de la densité a priori plus fréquemment.

2.3 Algorithmes quasi-statiques

Nous avons vu à la section précédente que la longueur efficace de la fenêtre d'assimilation est contrainte par la non quadraticité de la fonction de coût. Dans cette section nous revoyons et proposons des algorithmes permettant de dépasser ces problèmes de minimisation et d'atteindre de plus grandes fenêtres d'assimilation.

Les méthodes quasi-statiques ont été introduites par Pires et al. (1996) dans un contexte 4DVar. L'idée derrière le QSVA est de contrôler la manière d'ajouter les observations à la fonction de coût dans le but de maintenir le point de départ de chaque minimisation dans le bassin d'attraction du minimiseur global. La méthode consiste à répéter la minimisation avec un nombre croissant d'observations : la première minimisation est faite en n'utilisant que le premier vecteur d'observations, puis le nombre de vecteurs d'observations est augmenté et une nouvelle minimisation peut démarrer avec le minimiseur précédent comme point de départ. Le processus est ensuite répété jusqu'à ce que toutes les observations aient été prises en compte.

Cette procédure est directement applicable à la minimisation de la fonction de coût de l'IEnKS. Le schéma de gauche de la figure (2.6) schématise une minimisation QS et l'algorithme 2.1 est le pseudo-code d'un IEnKS avec minimisations quasi-statiques. Les nouveaux paramètres $(L_q)_{q < N_Q}$ contrôlent le nombre d'observations ajoutées à chaque minimisation, N_Q est le nombre total de minimisations intermédiaires.

La première ligne calcule la moyenne empirique et une matrice de déviation de l'ensemble a priori. La seconde applique le facteur d'inflation à la matrice de déviation. La

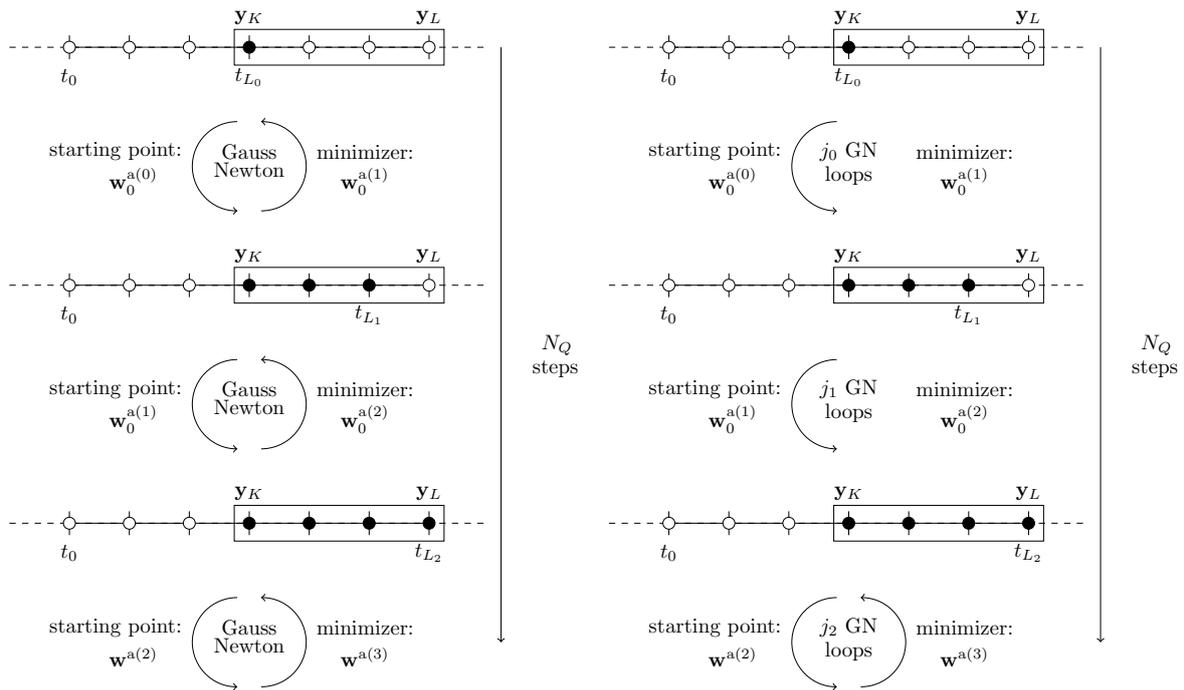


FIGURE 2.6 – Schémas des minimisations de l'IEnKS_{QS} (a) et de l'IEnKS_{QC} (b). Le rectangle contient les observations à assimiler et les points noirs représente les observations actives de la minimisation courante. Le « Gauss-Newton » encerclé de flèches représentent les itérations de la procédure. Le nombre d'étapes quasi-statiques est $N_Q = 3$. Le flux d'observations est contrôlé par les paramètres $(L_0, L_1, L_2) = (3, 5, 6)$. Pour l'IEnKS_{QC}, le nombre d'itérations de GN est contrôlé par les paramètres (j_0, j_1, j_2) .

Algorithm 2.1 Un cycle de l'EnKS_{QS/QC}

Require: Les paramètres de fenêtre $0 \leq K \leq L$, $S = L - K + 1$.

Les longueurs des fenêtres intermédiaires L_0, \dots, L_{N_Q-1} .

Le nombre d'itérations intermédiaires de Gauss-Newton j_0, \dots, j_{N_Q-1} .

Les observations $\mathbf{y}_{K:L}$.

L'ensemble a priori $\mathbf{E}_0^b \in \mathbb{R}^{n \times M}$.

La dimension de l'espace de l'ensemble $m = M - 1$.

Une base orthonormale $\begin{bmatrix} \frac{1}{\sqrt{M}} \mathbf{1}_M & \mathbf{U}_M \end{bmatrix} \in \mathbb{R}^{M \times M}$

Le pas des différences finies ϵ .

Les critères d'arrêt de Gauss-Newton Δ, j_{\max} .

L'inflation λ .

Ensure: \mathbf{E}_S^b is the next cycle prior ensemble

- 1: $\begin{bmatrix} \mathbf{x}_0^b & \mathbf{X}_0^b \end{bmatrix} = \mathbf{E}_0^b \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \end{bmatrix}$
- 2: $\mathbf{X}_0^b = \lambda \mathbf{X}_0^b$
- Analyse—
- 3: $\begin{bmatrix} \mathbf{w} & \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_m & \epsilon \mathbf{I}_m \end{bmatrix}$
- 4: **for** $q = 0, \dots, N_Q - 1$ **do**
- 5: $j = 0$
- 6: **repeat**
- 7: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{w} & \mathbf{I}_m \end{bmatrix}$
- 8: $\mathbf{E}_0 = \begin{bmatrix} \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w} & \mathbf{X}_0^b \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$
- 9: **for** $l = 0 \dots L_q$ **do**
- 10: **if** $l > 0$ **then**
- 11: $\mathbf{E}_l = \mathcal{M}_{l \leftarrow l-1}(\mathbf{E}_{l-1})$
- 12: **end if**
- 13: **if** $l \geq K$ **then**
- 14: $\begin{bmatrix} \mathbf{f} & \mathbf{F} \end{bmatrix} = \mathcal{H}_l(\mathbf{E}_l) \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \mathbf{W}^{-1} \end{bmatrix}$
- 15: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} += \begin{bmatrix} -\mathbf{F}^T \mathbf{R}_l^{-1} (\mathbf{y}_l - \mathbf{f}) & \mathbf{F}^T \mathbf{R}_l^{-1} \mathbf{F} \end{bmatrix}$
- 16: **end if**
- 17: **end for**
- 18: solve $\mathbf{C} \times \delta \mathbf{w} = -\mathbf{g}$ for $\delta \mathbf{w} \in \mathbb{R}^m$
- 19: $\mathbf{w} += \delta \mathbf{w}$
- 20: $j += 1$
- 21: **until** $\|\delta \mathbf{w}\| \leq \Delta$ or $j \geq j_{\max}/q$
- 22: **end for**
- Propagation—
- 23: solve $\mathbf{W} \mathbf{W}^T = \mathbf{C}^{-1}$ for $\mathbf{W} \in \mathbb{R}^{m \times m}$
- 24: $\mathbf{E}_0^a = \begin{bmatrix} \mathbf{x}_0^b + \mathbf{X}_0^b \mathbf{w} & \mathbf{X}_0^b \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$
- 25: $\mathbf{E}_S^a = \mathcal{M}_{S \leftarrow 0}(\mathbf{E}_0^a)$
- 26: $\mathbf{E}_S^b = \mathbf{E}_S^a$

ligne 3 initialise la moyenne et la matrice de déviation dans l'espace de l'ensemble. La boucle **for** lignes 4-22 répète les minimisations quasi-statiques. La ligne 5 initialise le compteur d'itérations de Gauss-Newton. La boucle **repeat** lignes 6-21 est la minimisation de Gauss-Newton. La ligne 7 initialise le gradient et la matrice hessienne de la fonction de coût. La ligne 8 engendre l'ensemble dans l'espace des états à t_0 à partir des statistiques dans l'espace de l'ensemble. La boucle **for** lignes 9-17 construit la fonction de coût contenant les observations jusqu'à t_{L_q} seulement. La ligne 11 propage d'un pas de temps l'ensemble courant. La ligne 14 linéarise l'opérateur \mathcal{F}_l à l'aide de l'ensemble. La ligne 15 utilise ces linéarisations pour mettre à jour le gradient et la hessienne de la fonction de coût. La ligne 18 résout le système linéaire de Gauss-Newton. La ligne 19 met à jour le minimum courant. La ligne 20 incrémente le compteur d'itérations. La ligne 21 sont les critères d'arrêt de Gauss-Newton. La ligne 23 met à jour l'anomalie de la variable d'ensemble avec une racine de l'inverse de la matrice hessienne. La ligne 24 engendre l'ensemble a posteriori de la variable d'état. La ligne 25 propage cet ensemble au temps t_S . La ligne 26 change cet ensemble a posteriori en ensemble a priori pour le prochain cycle. La figure 2.4 illustre ce schéma quasi-statique sur une seule analyse, comme expliqué dans la légende.

Pour cycler la méthode, la fenêtre est translatée de S pas de temps. Il est préférable de choisir S petit pour diminuer la déformation de la fonction de coût puisque, dans ce cas, peu de vecteurs d'observations entrent et sortent de la fenêtre d'assimilation. Cette nouvelle fonction de coût est ensuite minimisée avec la propagation du minimiseur précédent comme itéré initial. Bocquet and Sakov (2014) conseillent d'utiliser $S = 1$ pour n'assimiler qu'une seule observation à la fin de la fenêtre. C'est aussi la manière la plus simple d'assurer l'inégalité 2.34a. Cependant, il a été montré que cette méthode est sous-optimale à cause des fréquentes approximations gaussiennes de la densité a priori.

Une alternative est de garder $S = 1$ et d'utiliser toutes les observations de la fenêtre d'assimilation et donc de relâcher la condition $K = L - S + 1$. Les observations sont alors assimilées plusieurs fois. Ceci est fait dans la version MDA de l'IEnKS (Bocquet and Sakov, 2014). Pour que les statistiques restent cohérentes, au moins dans le cas linéaire et gaussien, les covariances d'erreur d'observation sont modifiées. Ainsi, le MDA IEnKS est un schéma quasi-statique et fait une bonne référence pour nos expériences numériques. Cependant, l'assimilation multiple des observations introduit des corrélations artificielles dans le cas non linéaire et non gaussien, ce qui conduit à des sous-optimalités. Une méthode semblable à l'IEnKS_{QS} a été utilisée avec succès dans Carrassi et al. (2017b) pour calculer la « model evidence ». Cependant, leur implémentation est basée sur la mise à jour de l'ensemble à chaque fois qu'un groupe d'observation est ajouté à la fonction

de coût, ce qui n'est pas aussi efficace numériquement que la méthode présentée ici.

L'efficacité des minimisations quasi-statiques repose sur le fait que, lorsqu'une observation est assimilée avec succès, la eMSE est réduite. Ainsi, la masse de probabilité se concentre et la moyenne a posteriori devient statistiquement plus proche de l'état. Il est alors plus probable que cette moyenne a posteriori se trouve dans un voisinage de l'état où le modèle est linéaire. La non quadraticité de la fonction de coût peut alors être augmentée en lui ajoutant de nouveaux termes d'observation. Cela est confirmé par l'argument suivant : notons $P^{(q)}$ la eMSE asymptotique de l'IEnKS à la q ème minimisation quasi-statique. Avec les notations et les suppositions de la section 2.2.3 et dans un contexte linéaire, on a la relation de récurrence suivante :

$$\left(P^{(q+1)}\right)^{-1} = \left(P^{(q)}\right)^{-1} + \Sigma_{L_{q+1}}^{L_{q+1}}, \quad (2.37a)$$

$$P^{(-1)} = n\alpha^{2S} \frac{\alpha^2 - 1}{\alpha^{2(L_{N_Q-1}+1)}}. \quad (2.37b)$$

Ainsi, $P^{(q+1)} < P^{(q)}$ et on peut augmenter la non quadraticité de la fonction de coût en lui ajoutant de nouveaux termes tant que la propagation des erreurs n'excède pas la borne :

$$\alpha^{2L_{q+1}} P^{(q)} \leq B. \quad (2.38)$$

Ce qui implique :

$$L_{q+1} \leq L_q + S_{\max}, \quad (2.39a)$$

$$L_0 \leq L_{N_Q-1} + S_{\max} - S, \quad (2.39b)$$

et finalement :

$$S \leq N_Q S_{\max}. \quad (2.40)$$

Les minimisations quasi-statiques permettent donc d'utiliser des fenêtres d'assimilations N_Q fois plus longues.

Malheureusement, ces minimisations quasi-statiques sont très coûteuses. En effet, elles ajoutent une troisième boucle externe qui répète N_Q fois la procédure de Gauss-Newton. Cependant, les minimisations intermédiaires sont inutilement précises. En effet, il suffit que les itérés initiaux soient dans le voisinage de $\mathbf{x}_{\infty S}$ où le modèle est pratiquement linéaire. On peut donc limiter le nombre de boucles de Gauss-Newton intermédiaires et laisser la convergence complète à la dernière minimisation. C'est ce qui est fait dans l'IEnKS quasi-convergent (IEnKS_{QC}) à l'aide des paramètres $(j_q)_{q < N_Q}$. Ils correspondent au nombre maximal d'itérations de Gauss-Newton à l'intérieur des

minimisations intermédiaires et sont typiquement égaux à 1 sauf pour la dernière minimisation QS. Le pseudo-code en rouge de l'algorithme 2.1 correspond à l'IEnKS_{QC} et le panneau de droite de la figure 2.6 en est un schéma.

2.4 Expériences numériques

Dans ce qui suit, des expériences numériques sont réalisées avec le modèle de Lorenz 1963 (L63) et de Lorenz 1995 (L95). L95 a été présentée à la section 2.2.4.1, L63 (Lorenz, 1963) est une modélisation simplifiée de la convection atmosphérique. Il est défini par l'équation différentielle ordinaire suivante :

$$\frac{dx}{dt} = \sigma (y - x), \quad (2.41a)$$

$$\frac{dy}{dt} = \rho x - y - xz, \quad (2.41b)$$

$$\frac{dz}{dt} = xy - \beta z. \quad (2.41c)$$

Ces équations sont intégrées avec un schéma de Runge-Kutta d'ordre 4, un pas de temps $\delta t = 0.01$ et $(\sigma, \rho, \beta) = (10, 28, 8/3)$. Avec ces paramètres, la dynamique de L63 est chaotique avec un plus grand exposant de Lyapunov proche de 0.91.

Ces deux modèles sont supposés parfaits. La trajectoire de l'état est engendrée depuis un point aléatoire de l'espace. L'ensemble initial est engendré par perturbation de l'état avec $\mathbf{B} = \mathbf{I}_n$ où $n = 40, 3$ et $m = 20, 3$ pour L95 et L63 respectivement. Les vecteurs d'observations sont aussi engendrés par perturbation de l'état avec $\mathcal{H} = \mathbf{R} = \mathbf{I}_n$ tous les $\Delta t = 0.05$ pour L95 et $\Delta t = 0.02$ pour L63. L'assimilation sur les $5 \times 10^3 \times \Delta t$ premier pas de temps est ignorée dans les deux cas.

Les paramètres de l'IEnKS sont $\epsilon = 10^{-4}$, $\delta = 10^{-3}$, $j_{\max} = 20$, $N_Q = 1$, $L_0 = L$. Pour l'IEnKS_{QS}, les paramètres QS sont $(L_0, L_1, \dots, L_{N_Q-1}) = K + (0, 1, \dots, N_Q - 1) \times \frac{S-1}{N_Q-1}$. Pour l'IEnKS_{QC}, les paramètres QS sont les mêmes et $(j_0, \dots, j_{N_Q-2}, j_{N_Q-1}) = (1, \dots, 1, 20)$. Les erreurs d'échantillonnage sont systématiquement prises en compte à l'aide des versions « finite-size » de l'IEnKS (Bocquet, 2011; Bocquet et al., 2015; Bocquet and Sakov, 2012) ce qui permet d'éviter le réglage coûteux de l'inflation. Finalement, la aRMSE est moyennée sur un nombre de cycle déterminé par le nombre d'observations assimilées. On utilise 5×10^5 vecteurs d'observation pour L95 et 5×10^6 vecteurs d'observation pour L63.

À la différence de Goodliff et al. (2015), nos expériences numériques ne concernent ni une non linéarité croissante, ni l'utilisation d'une matrice de covariance d'ébauche

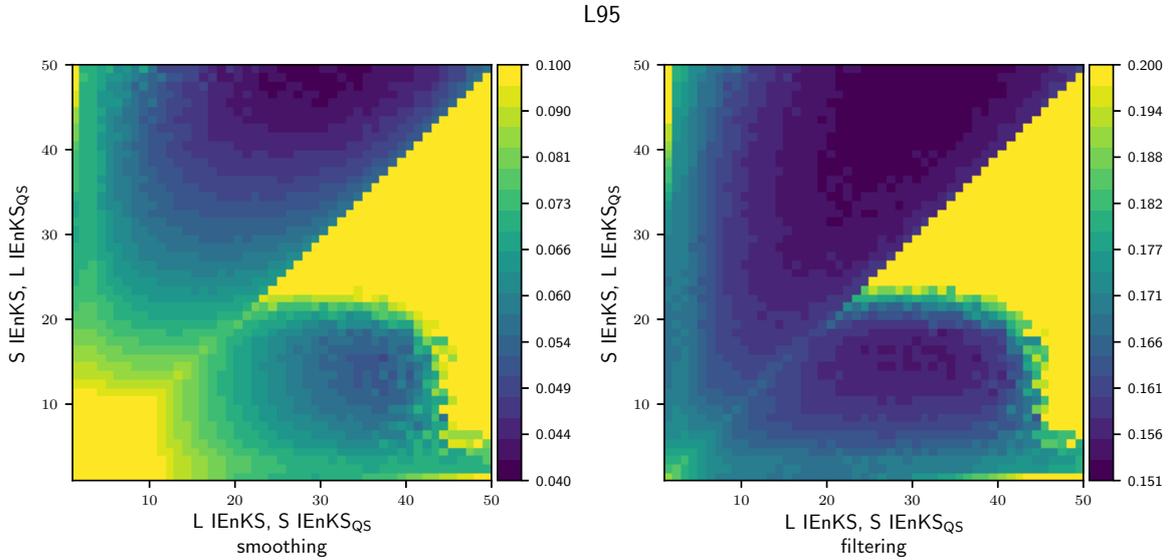


FIGURE 2.7 – aRMSEs de lissage et filtrage de l’IEnKS (triangles inférieurs) et de l’IEnKS_{QS} (triangles supérieurs, $N_Q = S$) en tant que fonction de L et S avec le modèle de L95. Le paramètre L est sur l’axe des abscisses pour l’IEnKS et sur l’axe des ordonnées pour l’IEnKS_{QS}. Le paramètre S est sur l’axe des abscisses pour l’IEnKS_{QS} et sur l’axe des ordonnées pour l’IEnKS. Les RMSEs de lissage au-dessus de 0.10 et de filtrage au-dessus de 0.20 sont de la même couleur pour une meilleure lisibilité. L’échelle est logarithmique.

climatologique. À la place, nous nous concentrons exclusivement sur la dépendance des performances de l’IEnKS avec les paramètres de fenêtre L, S ainsi que le nombre N_Q de minimisations QS. Goodliff et al. (2015) ont évalué numériquement l’approche QSVA avec des méthodes EnVar et hybrides, comme Bocquet and Sakov (2013), et confirment les idées de Pires et al. (1996) dans un contexte EnVar. En terme de précision, Goodliff et al. (2015) montrent que le lisseur de Kalman d’ensemble transformé (ETKS) surpasse toutes les méthodes hybrides de leurs expériences numériques. Puisque l’IEnKS surpasse l’ETKS (Bocquet and Sakov, 2013, 2014) tant que la longueur de la fenêtre d’assimilation n’est pas excessive, on en conclut que nos RMSEs seront plus petites ou égales à celles des méthodes hybrides reportées par Goodliff et al. (2015).

Les figures 2.7 et 2.8

affichent les aRMSEs de l’IEnKS et de l’IEnKS_{QS} pour les modèles L95 et L63 en fonction de S et L . Les performances de lissage et de filtrage de l’IEnKS augmentent pour de petites valeurs de L, S puis décroissent pour des valeurs plus grandes. Ceci est dû à l’apparence de minimums locaux. Comme remarqué par Goodliff et al. (2015) avec L63, la version QS des algorithmes permet d’atteindre de bien plus grandes fenêtres

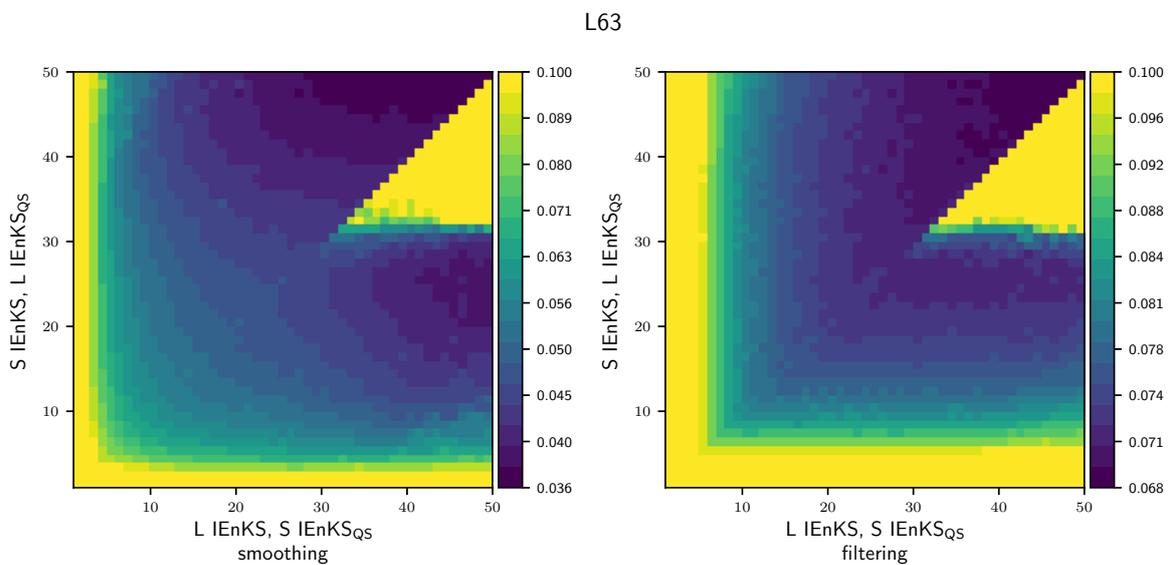


FIGURE 2.8 – aRMSEs de lissage et filtrage de l’IEnKS (triangles inférieurs) et de l’IEnKS_{QS} (triangles supérieurs, $N_Q = S$) en tant que fonction de L et S avec le modèle de L63. Le paramètre L est sur l’axe des abscisses pour l’IEnKS et sur l’axe des ordonnées pour l’IEnKS_{QS}. Le paramètre S est sur l’axe des abscisses pour l’IEnKS_{QS} et sur l’axe des ordonnées pour l’IEnKS. Les RMSEs de lissage au-dessus de 0.10 et de filtrage au-dessus de 0.20 sont de la même couleur pour une meilleure lisibilité. L’échelle est logarithmique.

d'assimilation et ainsi d'améliorer les performances. Cependant, les limites de la méthode deviennent visibles lorsque, pour le modèle L95 et $L = 50$, les meilleures aRMSEs sont atteintes pour $S < 50$. En revanche, les performances de filtrage de l'IEnKS_{QS} sont invariantes avec L et sont améliorées par S , comme pour les performances de filtrage du 4DVar dans la figure 2.2. Comme suggéré par Pires et al. (1996), on pourrait estimer la longueur utile de la fenêtre d'assimilation au-delà de laquelle le gain de performance est négligeable. Cependant, cette estimation se faisait pour la RMSE de filtrage d'un cycle du 4DVar. Cette estimation n'est pas directement pertinente pour un IEnKS_{QS} cyclé car, ici, une grande quantité d'observations ont déjà été assimilées et leur information condensée dans les statistiques d'ébauche. Ainsi, le gain de performance avec la longueur de fenêtre provient de la précision de l'approximation gaussienne de l'ébauche, précision ne pouvant pas être évaluée par la théorie linéaire. Par exemple, cette théorie linéaire prévoit à tort Eq. (2.25) l'indépendance des performances de filtrage ($l = L$) d'avec les paramètres de fenêtre.

La figure 2.9 compare les aRMSEs de lissage (première colonne), de filtrage (seconde colonne) et le nombre de propagations de l'ensemble (troisième colonne) de l'IEnKS_{QS} ($N_Q = S, S = L$), de l'IEnKS ($S = L$) et de l'IEnKS-MDA ($S = 1$) pour les modèles de L95 et L63 en fonction de L . Le nombre de propagations de l'ensemble est le nombre total de propagations de Δt divisé par le nombre total de vecteurs d'observations assimilés. Pour $L < 20$, les trois algorithmes affichent une amélioration des performances avec L . Pour $L > 20$, les performances de l'IEnKS se détériorent à cause des difficultés de minimisation. Pour $L < 40$, l'IEnKS_{QS} a de plus petites aRMSEs que l'IEnKS-MDA. Puisque l'IEnKS_{QS} est SDA par conception, il ne souffre pas des sous-optimalités liées à l'assimilation multiple en contexte non linéaire. De plus, l'IEnKS_{QS} utilise toujours moins de propagations de l'ensemble ; ce qui améliore le coût de calcul. Cependant, pour $L > 40$ et avec le modèle L95, l'approche QS ne permet plus de soutenir la non linéarité et les aRMSEs de l'IEnKS_{QS} se dégradent. Ainsi, l'IEnKS-MDA $S = 1$ a encore les meilleures performances. Avec le modèle de L63, l'IEnKS_{QS} est toujours meilleur que l'IEnKS-MDA, ce qui suggère que L peut encore être augmenté.

La figure 2.10 compare les aRMSEs de lissage (première colonne), de filtrage (seconde colonne) et le nombre de propagations de l'ensemble (troisième colonne) de l'IEnKS_{QS} et de l'IEnKS_{QC} en fonction du paramètre N_Q pour les modèles de L95 et L63. Les aRMSEs décroissent brutalement après un certain point pour les deux algorithmes et les deux modèles. Avant ce point, les algorithmes échouent à trouver le minimiseur global et la aRMSE est proche de la variance climatologique. Après ce point, les algorithmes trouvent le minimiseur global et les aRMSEs sont basses. Pour l'IEnKS_{QS} avec L95,

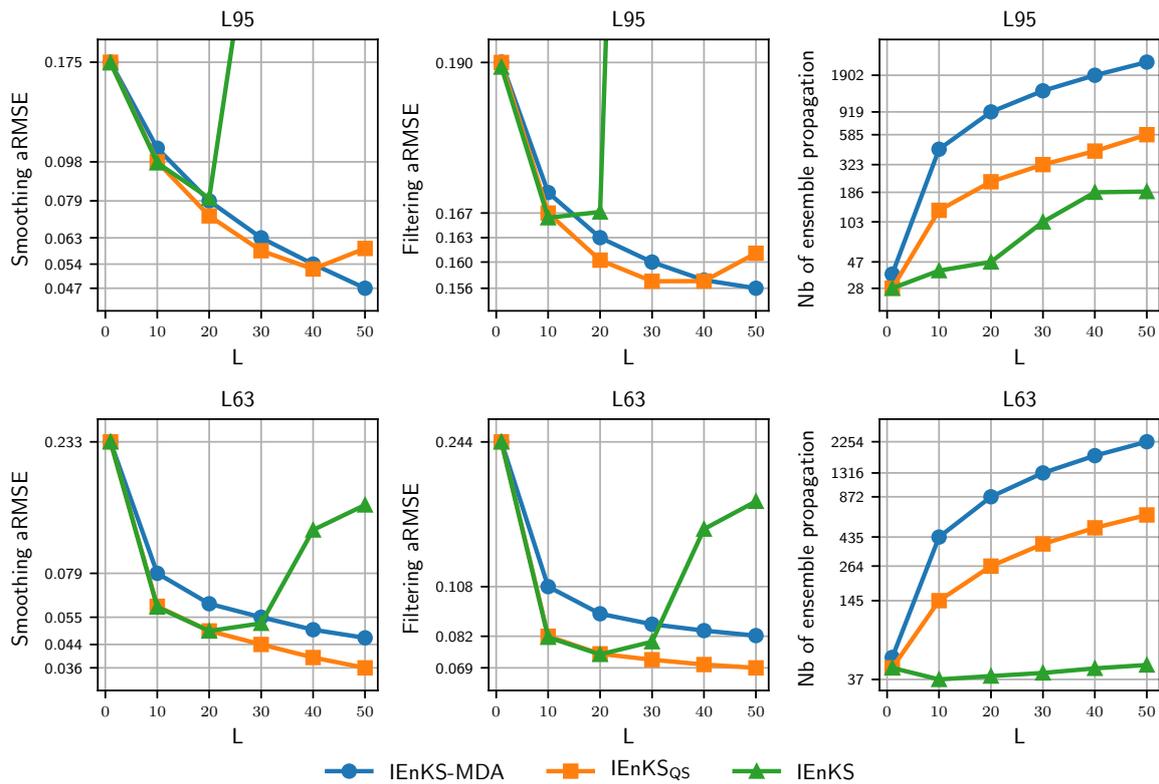


FIGURE 2.9 – aRMSEs de filtrage, de lissage et nombre de propagation de l'ensemble de l'IEnKS-MDA ($S=1$), l'IEnKS_{QS} et l'IEnKS ($S=L, N_Q=S$) en tant que fonction de L avec les modèles de L95 et L63. L'échelle est logarithmique et les propagations d'ensemble sont en multiple de Δt . Par exemple, avec $L=50$, une seule propagation de l'ensemble à la fin de la fenêtre d'assimilation compte pour 50.

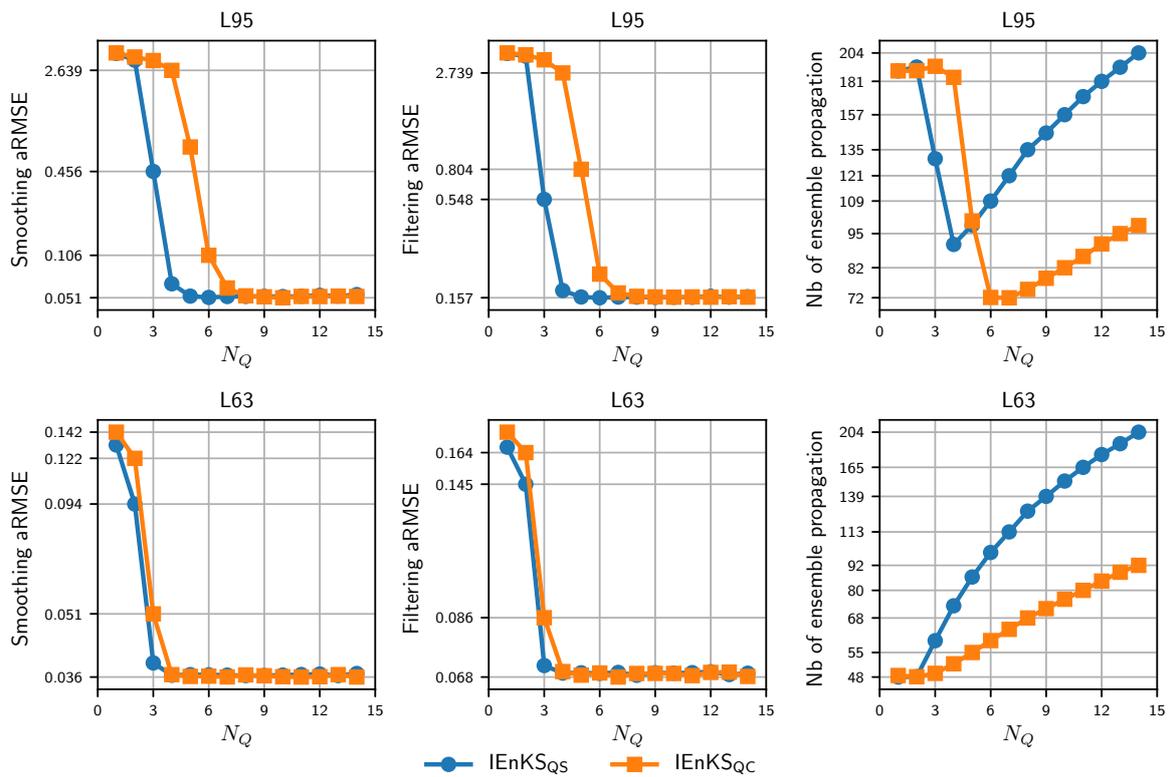


FIGURE 2.10 – aRMSEs de lissage, de filtrage et nombre de propagations de l'ensemble de l'IEnKS_{QS} et de l'IEnKS_{QC} ($S = L = 50$) en fonction de N_Q pour les modèles de L95 et L63. Le nombre de propagations de l'ensemble sont en multiple de Δt et l'échelle est logarithmique.

ce point peut être estimé à l'aide des résultats de la section 2.3 par $S/S_{\max} \simeq 3.6$. Ce qui correspond remarquablement à la Fig. 2.10. Ce point arrive plus tardivement pour l'IEnKS_{QC} mais il demande moins de propagations de l'ensemble, ce qui rend l'algorithme moins coûteux numériquement que l'IEnKS_{QS}. Cependant, ces propagations de l'ensemble ont un comportement différent pour L95 et L63 quand la minimisation échoue à trouver le minimum global. Pour le modèle de L95, le nombre de propagations de l'ensemble est élevé ce qui signifie que la minimisation demande beaucoup d'itérations et échoue à converger. Pour le modèle de L63, le nombre de propagations de l'ensemble est bas, ce qui signifie que la minimisation converge mais vers un extremum n'étant pas le minimiseur global.

2.5 Conclusion

Nous avons étendu dans ce papier l'étude de Pires et al. (1996) focalisée sur l'assimilation de données quasi-statique avec le 4DVar aux techniques séquentielles d'assimilation de données et plus spécifiquement aux méthodes 4D variationnelles d'ensemble, une référence du genre étant le lisseur de Kalman d'ensemble itératif (IEnKS).

L'impact du cyclage à long terme a été exploré théoriquement dans un contexte linéaire pour le 4DVar et l'IEnKS puis numériquement pour l'IEnKS dans un contexte non linéaire. La façon dont l'information est transportée entre les cycles d'assimilation explique la différence entre le 4DVar et l'IEnKS. Ces deux algorithmes révèlent une amélioration des performances avec le paramètre S qui compte le nombre de vecteurs d'observations assimilées lors du cycle aussi bien que le décalage temporel de la fenêtre d'assimilation entre les cycles. C'est une conséquence de l'approximation gaussienne de l'ébauche : plus grand est S , moins l'assimilation repose sur cette approximation.

Cependant, il a été observé que cette amélioration a une limite dans le cas d'un modèle parfait et chaotique. Le bassin d'attraction du minimiseur global de la fonction de coût se contracte avec L . A cause de cela, la méthode de Gauss-Newton échoue à trouver le minimiseur global de la fonction de coût, ce qui détériore les performances d'assimilation.

Les minimisations quasi-statiques mènent doucement mais sûrement au minimiseur global en répétant la minimisation de la fonction de coût. A mesure que la longueur de la fenêtre d'assimilation L augmente, les minimisations successives produisent des itérés initiaux restant dans le bassin d'attraction de l'argminimum global de la fonction de coût. Pour la plupart des couples S, L , l'IEnKS quasi-statique se révèle être une alternative plus précise et moins coûteuse que l'IEnKS avec assimilation multiple (IEnKS-MDA).

Malheureusement, l'IEEnKS ajoute une troisième boucle pouvant augmenter significativement son coût de calcul. Puisque la précision sur les minimiseurs intermédiaires est superflue, on peut limiter le nombre d'itérations des méthodes de Gauss-Newton intermédiaires.

Nous ne nous sommes pas focalisés sur l'applicabilité de ces méthodes en haute dimension ou avec des modèles imparfaits. En particulier, nous avons considéré de très longues fenêtres d'assimilation. Ce qui est intéressant d'un point de vue mathématique mais moins pertinent pour des modèles entachés d'erreurs. Cependant, nous savons de [Swanson et al. \(1998\)](#) que les résultats avec modèle parfait peuvent être étendus au cas imparfait lorsque le taux de croissance de l'erreur modèle est comparable au taux de croissance du plus gros vecteur de Lyapunov du modèle.

Chapitre 3

Le lisseur de Kalman itératif faiblement contraint

3.1 Introduction

Au cours des dernières décennies, les méthodes ensemble variationnelles (EnVar) sont devenues de plus en plus populaires (e.g., Chapitre 7 de [Asch et al., 2017](#)) dans le domaine de l'assimilation de données géophysiques. Ceci est expliqué par la synergie entre l'estimation des matrices de covariance d'erreurs et l'analyse variationnelle. D'un côté, la quantification de l'incertitude via un ensemble permet de cycler les étapes d'assimilation de manière naturelle. D'un autre côté, l'analyse variationnelle peut reposer sur des techniques itératives d'optimisation pour gérer la non linéarité des opérateurs.

3.1.1 Méthodes 4D ensemble variationnelles faiblement contraintes

Prendre en compte l'erreur modèle dans des algorithmes EnVar quadridimensionnels (4DEnVar) est difficile. En assimilation de données fortement contrainte, la trajectoire des états dans la fenêtre d'assimilation (DAW) est déterminée par sa condition initiale. Ceci n'est plus vrai en assimilation de donnée faiblement contrainte car le modèle est imparfait. Ainsi, la dimension de la variable de contrôle augmente considérablement avec la longueur de la fenêtre d'assimilation. Comme souligné par [Trémolet \(2006\)](#), ceci provoque des difficultés d'ordre numérique. Il y a donc un manque à gagner puisque la précision de l'assimilation s'améliore avec cette longueur de fenêtre ([Bocquet and Sakov, 2014](#); [Fillion et al., 2018](#)).

Une poignée de méthodes 4DEnVar faiblement contraintes ont été proposées dans la littérature. Dans [Mandel et al. \(2016\)](#), la fonction de coût 4DVar est minimisée suivant

les trajectoires de l'état avec une méthode itérative de résolution de sous problèmes linéarisés. Ces linéarisations sont obtenues par différence finie des opérateurs. Un lisseur de Kalman d'ensemble stochastique est utilisé comme minimiseur de la fonction de coût quadratique associée. L'itération de la méthode permet d'améliorer les performances. Cependant, les perturbations stochastiques des observations et de la trajectoire introduisent des erreurs d'échantillonnage. Dans [Amezcuca et al. \(2017\)](#), le forçage d'erreur modèle est supposé gaussien alors que le biais d'erreur modèle est la variable de contrôle. Cette variable de contrôle est alors traitée dans leur fonction de coût 4DVar comme si elle était gaussienne et d'autocovariance nulle. De telles approximations sont faites pour éviter certains termes de la fonction de coût nécessitant l'évaluation du modèle. Il en résulte que leur fonction de coût est une fonction quadratique du biais d'erreur modèle. Nous montrerons dans la section [3.2.3](#) que la minimisation de cette fonction de coût peut être découplée et ramenée à une forme fortement contrainte.

Le filtre de Kalman d'ensemble itératif (IEnKF-Q) de [Sakov et al. \(2018\)](#) est un autre type de méthode EnVar prenant en compte l'erreur modèle additive. Il utilise une minimisation variationnelle dans le sous-espace de l'ensemble sur une fenêtre d'assimilation d'un pas de temps. Les dérivées sont estimées à l'aide des différences finies et l'ensemble est mis à jour de manière déterministe. L'IEnKF-Q est la généralisation naturelle du filtre de Kalman d'ensemble itératif (IEnKF, [Sakov et al., 2012](#)) au cas faiblement contraint. Une autre généralisation de l'IEnKF tenant compte de fenêtres d'assimilation plus générales est le lisseur de Kalman d'ensemble itératif (IEnKS, [Bocquet and Sakov, 2014](#)). Il coïncide avec l'IEnKF lorsque la fenêtre d'assimilation s'étend sur un seul pas de temps. De plus longues fenêtres d'assimilation améliorent les performances de lissage et de filtrage avec des modèles non linéaires puisqu'elles réduisent la fréquence de l'approximation gaussienne des densités a priori ([Fillion et al., 2018](#)).

Le but de ce papier est de construire algorithmiquement et de tester numériquement un lisseur de Kalman d'ensemble itératif faiblement contraint (IEnKS-Q) qui sera conçu pour être la généralisation faiblement contrainte de l'IEnKS, aussi bien que la généralisation de l'IEnKF-Q à de plus grandes fenêtres d'assimilation. Un premier pas dans cette direction a été accompli par [Sakov and Bocquet \(2018\)](#) où l'algorithme d'une itération de l'IEnKS-Q a été donnée.

3.1.2 Plan

La section [3.2](#) est un rappel des concepts centraux nécessaires à la construction de l'IEnKS-Q. Les connections avec la méthode 4DEnVar faiblement contrainte de [Amezcuca et al. \(2017\)](#) sont ensuite clarifiées. La section [3.3](#) développe la construction théorique de

de l'IEEnKS-Q, i.e. une méthode 4DEnVar faiblement contrainte déterministe. La complexité numérique de l'IEEnKS-Q est aussi étudiée et des solutions faisant potentiellement face au nombre élevé d'évaluations du modèle sont données. Des algorithmes de l'IEEnKS-Q sont proposés dans la section 3.4. Ils sont numériquement testés dans la section 3.5 avec le modèle chaotique de Lorenz à 40 variables (Lorenz and Emanuel, 1998). Les conclusions sont tirées dans la section 3.6.

3.2 Assimilation de données séquentielle faiblement contrainte dans l'espace de l'état

Cette section définit le système séquentiel d'assimilation de données avec erreur modèle additive considérée dans ce papier. Il est basé sur une fenêtre d'assimilation se déplaçant dans le temps dont les paramètres sont définis. La fonction de coût de l'analyse de ce système est établie sans référence aux ensembles à cette étape. Ensuite, il est montré que l'approche suggérée par Amezcua et al. (2017) peut être simplifiée suivant une extension de la preuve de Sakov et al. (2018) dans le cas de fenêtres d'un pas de temps.

3.2.1 Le système d'assimilation de données

Soient $(\mathbf{x}_l)_{l \in \mathbb{Z}}$ et $(\mathbf{y}_l)_{l \in \mathbb{Z}}$ les vecteurs aléatoires avec valeurs dans \mathbb{R}^n et \mathbb{R}^d définies au temps $(t_l)_{l \in \mathbb{Z}}$. Ils représentent l'état du système et les observations de ce système. Ils évoluent et sont reliés par les équations :

$$\mathbf{x}_l = \mathcal{M}_l(\mathbf{x}_{l-1}) + \boldsymbol{\varepsilon}_l^q, \quad (3.1a)$$

$$\mathbf{y}_l = \mathcal{H}_l(\mathbf{x}_l) + \boldsymbol{\varepsilon}_l^r, \quad (3.1b)$$

pour $l \in \mathbb{Z}$, où $(\mathcal{M}_l)_{l \in \mathbb{Z}}$ est le modèle dynamique et $(\mathcal{H}_l)_{l \in \mathbb{Z}}$ les opérateurs d'observation. Les forçages d'erreur modèle $(\boldsymbol{\varepsilon}_l^q)_{l \in \mathbb{Z}}$ et les erreurs d'observation $(\boldsymbol{\varepsilon}_l^r)_{l \in \mathbb{Z}}$ sont supposées mutuellement indépendantes et gaussiennes :

$$\begin{aligned} \boldsymbol{\varepsilon}_l^q &\sim \mathcal{N}(\mathbf{0}_n, \mathbf{Q}_l), \\ \boldsymbol{\varepsilon}_l^r &\sim \mathcal{N}(\mathbf{0}_d, \mathbf{R}_l). \end{aligned} \quad (3.2a)$$

Les matrices $\mathbf{Q}_l \in \mathbb{R}^{n \times n}$ et $\mathbf{R}_l \in \mathbb{R}^{d \times d}$ sont les matrices de covariance de l'erreur modèle et de l'erreur d'observation respectivement. Dans ce qui suit, une notation 4D est utilisée

pour concaténer verticalement les vecteurs et les ensembles suivant le temps :

$$\mathbf{x}_{l_1:l_2} = \text{vec}(\mathbf{x}_{l_1}, \dots, \mathbf{x}_{l_2}) = \begin{bmatrix} \mathbf{x}_{l_1} \\ \vdots \\ \mathbf{x}_{l_2} \end{bmatrix}, \quad (3.3a)$$

$$\mathbf{E}_{l_1:l_2} = \text{vec}(\mathbf{E}_{l_1}, \dots, \mathbf{E}_{l_2}) = \begin{bmatrix} \mathbf{E}_{l_1} \\ \vdots \\ \mathbf{E}_{l_2} \end{bmatrix}, \quad (3.3b)$$

pour $l_1 \leq l_2$. La notation 4D d'une suite de matrices de covariance est bloc diagonale quand le vecteur 4D associé est décorrélé suivant le temps :

$$\mathbf{B}_{l_1:l_2} = \text{diag}(\mathbf{B}_{l_1}, \dots, \mathbf{B}_{l_2}) = \begin{bmatrix} \mathbf{B}_{l_1} & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \mathbf{B}_{l_2} \end{bmatrix}. \quad (3.4)$$

L'indice « : l_1 » indique que tous les pas de temps depuis le début de l'assimilation sont inclus jusqu'à l_1 .

3.2.2 Principes d'assimilation de données séquentielle

Notre objectif est d'estimer séquentiellement la densité a posteriori :

$$p(\mathbf{x}_{0:L} | \mathbf{y}_{:L}), \quad (3.5)$$

de la trajectoire de l'état connaissant toutes les observations passées. Quitte à translater les indices temporels, nous pouvons choisir de faire démarrer la fenêtre d'assimilation courante à t_0 . Ainsi, le paramètre $L \geq 0$ est à la fois un indice relatif de longueur de fenêtre et l'indice temporel absolu de la fin de la fenêtre courante. L'assimilation de l'ensemble des observations passées est divisée en cycles. Un décalage de la fenêtre d'assimilation de S pas de temps assure la transition entre ces cycles. Ainsi, la densité a priori du cycle courant est $p(\mathbf{x}_{0:L} | \mathbf{y}_{:L-S})$ alors que $p(\mathbf{x}_{S:L+S} | \mathbf{y}_{:L})$ est la densité a priori du prochain cycle.

Soit $K \leq L$ l'indice temporel tel que $\mathbf{y}_{K:L}$ est le nouveau paquet de vecteurs d'observations à assimiler lors de ce cycle. Si les observations ne sont assimilées qu'une seule et unique fois, les paramètres de fenêtres doivent vérifier :

$$L - S = K - 1. \quad (3.6)$$

Les densités a priori et a posteriori du cycle courant sont reliées à la densité a priori du cycle suivant par les équations suivantes :

$$p(\mathbf{x}_{0:L}|\mathbf{y}_{:L}) \propto p(\mathbf{y}_{K:L}|\mathbf{x}_{K:L})p(\mathbf{x}_{0:L}|\mathbf{y}_{:L-S}) \quad [\text{Analyse}], \quad (3.7a)$$

$$p(\mathbf{x}_{S:L+S}|\mathbf{y}_{:L}) \propto \int p(\mathbf{x}_{L+1:L+S}|\mathbf{x}_L)p(\mathbf{x}_{0:L}|\mathbf{y}_{:L})d\mathbf{x}_{0:S-1} \quad [\text{Propagation}], \quad (3.7b)$$

où le symbole \propto signifie « égal à une constante multiplicative près » lorsqu'il est appliqué à des densités. Les équations Eq. (3.7) découlent de la règle de Bayes et de l'indépendance des erreurs du système Eq. (3.1). Elles sont associées à l'étape d'analyse et de propagation.

Pour que l'assimilation soit réalisable, une approximation gaussienne de la densité a priori est faite lors de l'étape d'analyse. Cependant, cette approximation peut ne pas être adaptée à la trajectoire $\mathbf{x}_{0:L}$. L'hypothèse Eq. (3.2a) nous suggère de changer les variables d'état \mathbf{x}_l pour les forçages d'erreur modèle $\boldsymbol{\varepsilon}_l^q$ pour éviter la non linéarité du modèle. Nous considérerons donc le changement de variable suivant, rappelant les transformations suggérées par Trémolet (2006) :

$$\mathcal{G} : \mathbf{x}_{0:L} \mapsto \mathbf{v}_{0:L} = \begin{bmatrix} \mathbf{x}_0 \\ \vdots \\ \mathbf{x}_G \\ \mathbf{x}_{G+1} - \mathcal{M}_{G+1}(\mathbf{x}_G) \\ \vdots \\ \mathbf{x}_L - \mathcal{M}_L(\mathbf{x}_{L-1}) \end{bmatrix}, \quad (3.8)$$

le nouveau paramètre de fenêtre d'assimilation $G \in \{0, \dots, L\}$ est l'indice temporel de la dernière variable d'état inchangée. La variable $\mathbf{v}_{0:L} = \text{vec}(\mathbf{x}_{0:G}, \boldsymbol{\varepsilon}_{G+1:L}^q)$ est la nouvelle variable d'ébauche. Le fait que le paramètre G influe sur le caractère gaussien de cette variable d'ébauche est illustré par l'exemple suivant dans la fenêtre d'assimilation $L = 1$ avec les hypothèses $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{x}_0^b, \mathbf{B}_0)$ et $\boldsymbol{\varepsilon}_1^q \sim \mathcal{N}(\mathbf{0}_n, \mathbf{Q}_1)$. Avec $G = 0$, la variable $\mathbf{v}_{0:1} = \text{vec}(\mathbf{x}_0, \boldsymbol{\varepsilon}_1^q)$ est gaussienne ; alors qu'avec $G = 1$, la variable $\mathbf{v}_{0:1} = \text{vec}(\mathbf{x}_0, \mathbf{x}_1)$ ne l'est pas à cause de la non linéarité du modèle dans $\mathbf{x}_1 = \mathcal{M}_1(\mathbf{x}_0) + \boldsymbol{\varepsilon}_1^q$. Le changement de variable \mathcal{G} Eq. (3.8) est triangulaire ou causal puisque chaque ligne l est la somme de l'état présent \mathbf{x}_l et d'une fonction des états passés $\mathbf{x}_{0:l-1}$. Clairement, \mathcal{G} est inversible et le déterminant de sa matrice jacobienne est 1. Ainsi, l'approximation gaussienne de la densité de la variable d'ébauche $\mathbf{v}_{0:L}$ est :

$$p(\mathbf{v}_{0:L}|\mathbf{y}_{:L-S}) \simeq \mathcal{N}(\mathbf{v}_{0:L}^b, \mathbf{B}_{0:L}), \quad (3.9)$$

où $\mathcal{N}(\mathbf{v}_{0:L}^b, \mathbf{B}_{0:L})$ est la densité gaussienne de moyenne $\mathbf{v}_{0:L}^b$ et de matrice de covariance $\mathbf{B}_{0:L}$. Ces statistiques sont fournies par le cycle d'assimilation précédent. Une représentation graphique de la fenêtre d'assimilation et de ses paramètres est donnée Fig. (3.1).

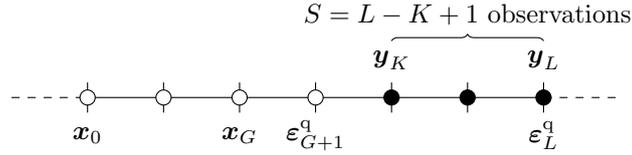


FIGURE 3.1 – Une fenêtre d'assimilation paramétrisée par $G, K, L, S = L - K + 1$. Les points blancs sont des observations déjà assimilées ; les points noirs sont les nouvelles observations.

L'approximation gaussienne de la densité a priori Eq. (3.9) est utilisée dans l'opposé du logarithme de Eq. (3.7a) avec les notations 4D :

$$\mathbf{x}, \mathbf{y}, \mathbf{v} = \mathbf{x}_{0:L}, \mathbf{y}_{K:L}, \mathbf{v}_{0:L}, \quad (3.10a)$$

$$\mathbf{v}^b, \mathbf{B}, \mathbf{R} = \mathbf{v}_{0:L}^b, \mathbf{B}_{0:L}, \text{diag}(\mathbf{R}_K, \dots, \mathbf{R}_L), \quad (3.10b)$$

$$\mathcal{H} : \mathbf{x} \mapsto \begin{bmatrix} \mathcal{H}_K(\mathbf{x}_K) \\ \vdots \\ \mathcal{H}_L(\mathbf{x}_L) \end{bmatrix}, \quad (3.10c)$$

pour exprimer de façon compacte la fonction de coût 4DVar faiblement contrainte de la variable \mathbf{x} ou \mathbf{v} :

$$J(\mathbf{x}; \mathbf{y}) \propto \frac{1}{2} \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathcal{G}(\mathbf{x}) - \mathbf{v}^b\|_{\mathbf{B}^{-1}}^2, \quad (3.11a)$$

$$J(\mathbf{v}; \mathbf{y}) \propto \frac{1}{2} \|\mathbf{y} - \mathcal{H} \circ \mathcal{G}^{-1}(\mathbf{v})\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\mathbf{v} - \mathbf{v}^b\|_{\mathbf{B}^{-1}}^2, \quad (3.11b)$$

où $\|\mathbf{z}\|_{\mathbf{S}}^2 = \mathbf{z}^T \mathbf{S} \mathbf{z}$ et le symbole \propto signifie « égal à une constante additive près » lorsqu'il est appliqué à des fonctions de coût. Notons que plus G est grand, moins le modèle est évalué lors du calcul de la fonction de coût. La minimisation variationnelle de la fonction de coût Eq. (3.11b) conduit à l'approximation quadratique :

$$J(\mathbf{v}; \mathbf{y}) \simeq \frac{1}{2} \|\mathbf{v} - \mathbf{v}^a\|_{(\mathbf{C}^a)^{-1}}^2, \quad (3.12)$$

dont l'opposé de l'exponentielle est proportionnel à une approximation gaussienne de la densité a posteriori :

$$p(\mathbf{v}|\mathbf{y}_{:L}) \simeq \mathcal{N}(\mathbf{v}^a, \mathbf{C}^a). \quad (3.13)$$

L'approximation gaussienne de la densité a posteriori est appliquée à la variable \mathbf{v} plutôt que \mathbf{x} pour les mêmes raisons que l'approximation gaussienne de la densité a priori.

Pendant l'étape de propagation, la variable d'ébauche \mathbf{v} est étendue avec les forçages d'erreur modèle pour former le vecteur aléatoire $\text{vec}(\mathbf{x}_{0:G}, \boldsymbol{\varepsilon}_{G+1:L+S}^q)$ dont la moyenne et matrice de covariance connaissant $\mathbf{y}_{:L}$ sont :

$$\text{vec}(\mathbf{v}^a, \mathbf{0}_n, \dots, \mathbf{0}_n), \quad (3.14a)$$

$$\text{diag}(\mathbf{C}^a, \mathbf{Q}_{L+1}, \dots, \mathbf{Q}_{L+S}), \quad (3.14b)$$

grâce à l'indépendance des erreurs. Le modèle est ensuite utilisé pour propager ces statistiques afin d'obtenir celles du vecteur aléatoire $\text{vec}(\mathbf{x}_{0:G+S}, \boldsymbol{\varepsilon}_{G+S+1:L+S}^q)$ où l'indice G a été déplacé de S pas de temps dans le futur. Ces statistiques sont alors réduites pour obtenir la moyenne \mathbf{v}^b et la matrice de covariance \mathbf{B}' de la variable d'ébauche $\mathbf{v}' = \text{vec}(\mathbf{x}_{S:G+S}, \boldsymbol{\varepsilon}_{G+S+1:L+S}^q)$.

Les développements précédents généralisent l'assimilation de données fortement contrainte avec une fenêtre d'assimilation se déplaçant dans le temps développée dans e.g. [Bocquet and Sakov \(2014\)](#); [Fillion et al. \(2018\)](#) à un cadre faiblement contraint. Un cycle de la méthode est illustré Fig. (3.2)

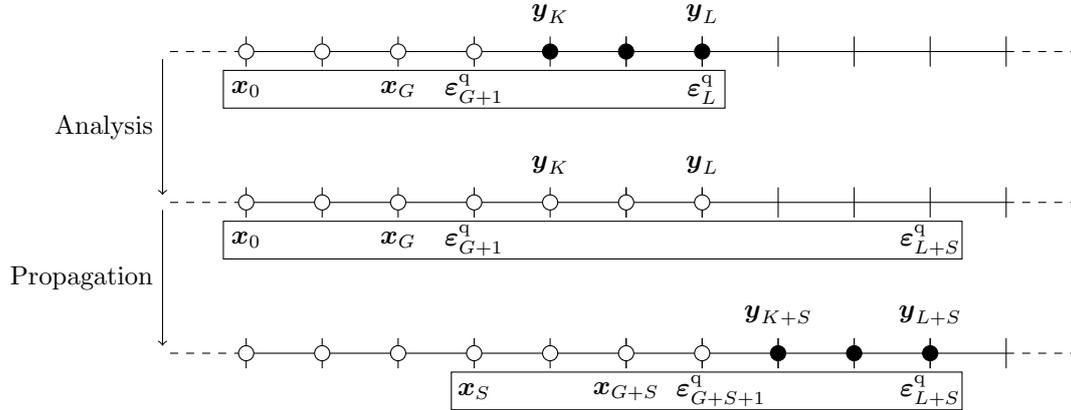


FIGURE 3.2 – Un cycle d’assimilation faiblement contrainte. Pendant l’analyse, les statistiques a priori de $G + 1$ états $\mathbf{x}_0, \dots, \mathbf{x}_G$ et $L - G$ forçage d’erreur modèle $\boldsymbol{\varepsilon}_{G+1}^q, \dots, \boldsymbol{\varepsilon}_L^q$ (premier rectangle) sont transformées en statistiques a posteriori par l’assimilation de S nouveaux vecteurs d’observations $\mathbf{y}_K, \dots, \mathbf{y}_L$ (les points noirs sont les nouvelles observations, alors que les points blancs sont les observations déjà assimilées). Pendant la propagation, ces statistiques a posteriori sont étendues avec les statistiques d’erreur modèle pour obtenir celles de $\mathbf{x}_0, \dots, \mathbf{x}_G, \boldsymbol{\varepsilon}_{G+1}^q, \dots, \boldsymbol{\varepsilon}_{L+S}^q$ (second rectangle). Une propagation de S pas de temps permet ensuite d’obtenir les statistiques de $\mathbf{x}_0, \dots, \mathbf{x}_{G+S}, \boldsymbol{\varepsilon}_{G+S+1}^q, \dots, \boldsymbol{\varepsilon}_{L+S}^q$. Elles sont finalement réduites pour obtenir les statistiques a priori de $\mathbf{x}_S, \dots, \mathbf{x}_G, \boldsymbol{\varepsilon}_{G+1}^q, \dots, \boldsymbol{\varepsilon}_{L+S}^q$ (troisième rectangle).

3.2.3 Comparaison avec le biais d’erreur modèle comme variable de contrôle

Suivant une ancienne suggestion de Trémolet (2006), les auteurs de Amezcua et al. (2017) ont choisis la première variable d’état \mathbf{x}_0 avec les variables de biais d’erreur modèle :

$$\boldsymbol{\beta}_l = \mathbf{x}_l - \mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0), \quad (3.15)$$

pour $l = 1, \dots, L$ comme variable de contrôle de leur fonction de coût 4DVar faiblement contrainte :

$$\begin{aligned} J(\mathbf{x}_0, \boldsymbol{\beta}_{1:L}; \mathbf{y}_{1:L}) &= \frac{1}{2} \sum_{l=1}^L \|\mathbf{y}_l - \mathcal{H}_l(\mathcal{M}_{l \leftarrow 0}(\mathbf{x}_0) + \boldsymbol{\beta}_l)\|_{\mathbf{R}_l}^2 \\ &\quad + \frac{1}{2} \sum_{l=1}^L \|\boldsymbol{\beta}_l\|_{\mathbf{Q}_l}^2 + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0}^2. \end{aligned} \quad (3.16)$$

Pour être bayésienne, c'est à dire pour être l'opposé du logarithme de la densité a posteriori $p(\mathbf{x}_0, \boldsymbol{\beta}_{1:L} | \mathbf{y}_{:L})$, cette fonction de coût requiert les hypothèses suivantes :

$$p(\mathbf{x}_0 | \mathbf{y}_{:0}) = \mathcal{N}(\mathbf{x}_0^b, \mathbf{B}_0), \quad (3.17a)$$

$$p(\boldsymbol{\beta}_l | \mathbf{y}_{:0}) = \mathcal{N}(\mathbf{0}_n, \mathbf{Q}_l^\beta), \quad (3.17b)$$

pour $l = 1, \dots, L$ et l'indépendance mutuelle sachant $\mathbf{y}_{:0}$ de la famille $(\mathbf{x}_0, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_L)$. L'exposant β distingue la covariance de biais d'erreur modèle de la covariance de forçage d'erreur modèle. Alors que pour être l'opposé du logarithme de $p(\mathbf{v}_{0:L} | \mathbf{y}_{:0})$ dans le cas $K = 1$, notre fonction de coût Eq. (3.11b) requiert :

$$p(\mathbf{x}_{0:G} | \mathbf{y}_{:0}) = \mathcal{N}(\mathbf{x}_{0:G}^b, \mathbf{B}_{0:G}), \quad (3.18a)$$

$$p(\boldsymbol{\varepsilon}_l^q | \mathbf{y}_{0:}) = \mathcal{N}(\mathbf{0}_n, \mathbf{Q}_l), \quad (3.18b)$$

pour $l = G+1, \dots, L$ et l'indépendance mutuelle sachant $\mathbf{y}_{:0}$ de la famille $(\mathbf{x}_{0:G}, \boldsymbol{\varepsilon}_{G+1}^q, \dots, \boldsymbol{\varepsilon}_L^q)$. L'annexe C.1 montre que ces deux ensembles d'hypothèses Eq. 3.17, 3.18 ne sont pas compatibles lorsque le modèle est non linéaire à moins que $G = 0$ et $L = 1$. L'approche de Amezcua et al. (2017) n'est donc pas généralement équivalente à celle de cette étude.

Cependant, dans cette étude comme dans Amezcua et al. (2017), des suppositions spécifiques sont faites sur les statistiques du système d'assimilation :

$$p(\boldsymbol{\varepsilon}_l^q) = \mathcal{N}(\mathbf{0}_n, \mathbf{Q}_l), \quad (3.19a)$$

$$p(\boldsymbol{\varepsilon}_l^r) = \mathcal{N}(\mathbf{0}_n, \mathbf{R}_l), \quad (3.19b)$$

pour tout l et toutes ces erreurs sont mutuellement indépendantes. Dans un contexte d'assimilation séquentielle avec un modèle non linéaire, les modélisations statistiques Eq. (3.17, 3.18) sont des approximations gaussiennes faites pour que l'assimilation soit calculable. Puisque ces deux modélisations sont incompatibles, la question est de déterminer quelle est celle qui approche le mieux les statistiques exactes du système d'assimilation.

D'abord, la modélisation statistique Eq. (3.17) réalise une approximation gaussienne stricte à chaque pas de temps t_0, \dots, t_L de la fenêtre d'assimilation; alors que la modélisation statistique Eq. (3.18) réalise une approximation gaussienne stricte aux pas de temps t_0, \dots, t_G seuls. En effet, les forçages d'erreur modèle $(\boldsymbol{\varepsilon}_l^q)_{G < l \leq L}$ sont indépendants de $\mathbf{y}_{:0}$ (puisque c'est une fonction de $\boldsymbol{\varepsilon}_{:0}^q, \boldsymbol{\varepsilon}_{:0}^r$ et $G \geq 0$) donc gaussiens d'après les suppositions Eq. (3.19). L'approximation gaussienne sur l'intégralité de la fenêtre d'assimilation de Eq. (3.17) simplifie l'interaction entre le modèle non linéaire et les forçages d'erreur modèle. Par conséquent, la fonction de coût Eq. (3.16) est quadratique selon

$\beta_{1:L}$ lorsque l'opérateur d'observation est linéaire. Il est ainsi montré dans l'annexe C.2 que la minimisation de la fonction de coût Eq. (3.16) se réduit à la minimisation d'une fonction de coût d'un IEnKS fortement contraint selon \mathbf{x}_0 . Le minimiseur selon $\beta_{1:L}$ est alors directement donné par la solution d'un système linéaire impliquant le minimiseur précédent. Ensuite, les covariances croisées entre les $(\beta_l)_{0 < l \leq L}$ sont négligées dans Eq. (3.17); alors que les covariances croisées entre les $(\varepsilon_l^q)_{G < l \leq L}$ sont nulles d'après les suppositions Eq. (3.19).

3.3 Construction de l'IEnKS-Q

La section 3.2 concernait le formalisme variationnel faiblement contraint dans l'espace de l'état. Dans ce formalisme, la fonction de coût Eq. (3.11b) est bien définie grâce à la supposée symétrie définie-positivité de $\mathbf{B}_{0:L}$. Cette propriété peut-être perdue lorsque les matrices de covariance sont estimées à l'aide d'ensembles. Dans ce cas, la variable de contrôle est encore changée pour une variable d'ensemble réduite.

3.3.1 Représentation dans le sous-espace de l'ensemble

Supposons qu'un ensemble $\mathbf{E}^b \in \mathbb{R}^{n(L+1) \times M}$ de M membres échantillonnant $p(\mathbf{v}|\mathbf{y}_{:L-S})$ Eq. (3.9) est disponible par récurrence. Ici la moyenne \mathbf{v}^b et la matrice de covariance \mathbf{B} sont définies par leurs estimateurs :

$$\mathbb{E}_{p(\mathbf{v}|\mathbf{y}_{:L-S})}[\mathbf{v}] \simeq \mathbf{v}^b = \mathbf{E}^b \frac{\mathbf{1}_M}{M}, \quad (3.20a)$$

$$\mathbb{C}_{p(\mathbf{v}|\mathbf{y}_{:L-S})}[\mathbf{v}] \simeq \mathbf{B} = (\mathbf{A}^b) (\mathbf{A}^b)^\top, \quad (3.20b)$$

$$\mathbf{A}^b = \mathbf{E}^b \frac{\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^\top}{M}}{\sqrt{M-1}}. \quad (3.20c)$$

Le vecteur $\mathbf{1}_M$ est le vecteur colonne de dimension M ne contenant que des 1, \mathbf{I}_M est la matrice identité dans $\mathbb{R}^{M \times M}$ et $\mathbf{A}^b \in \mathbb{R}^{n(L+1) \times M}$ est la matrice d'anomalies normalisées de \mathbf{E}^b . Cette matrice d'anomalies n'est pas injective puisque $\mathbf{1}_M$ est dans son noyau. Il en résulte un changement de variable mal défini dans l'espace de l'ensemble. Ainsi, un degré de liberté doit être fixé à l'aide d'un terme de jauge dans Bocquet and Sakov (2014). Ici, le problème est évité en utilisant un facteur injectif \mathbf{V}^b de \mathbf{B} tel que :

$$\mathbf{B} = (\mathbf{V}^b) (\mathbf{V}^b)^\top. \quad (3.21)$$

Un tel facteur injectif \mathbf{V}^b de la matrice de covariance empirique de \mathbf{E}^b sera appelé « une matrice de déviation de l'ensemble \mathbf{E}^b » par opposition à sa matrice d'anomalies.

Pour calculer une matrice de déviation, soit :

$$\left[\begin{array}{c|c} \frac{\mathbf{1}_M}{M} & \mathbf{U}_M \end{array} \right] \in \mathbb{R}^{M \times M}, \quad (3.22)$$

une matrice orthogonale dont la première colonne est la normalisation de $\mathbf{1}_M$ et le reste est la matrice $\mathbf{U}_M \in \mathbb{R}^{M \times (M-1)}$. Une telle matrice est, par exemple, la matrice de Householder (e.g. [Stoer and Bulirsch, 1982](#)) envoyant $\frac{\mathbf{1}_M}{\sqrt{M}}$ sur le premier vecteur de la base canonique. Puisque cette matrice est orthogonale, nous avons :

$$\mathbf{U}_M \mathbf{U}_M^T = \mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M}, \quad (3.23a)$$

$$= \left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right) \left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right)^T, \quad (3.23b)$$

par symétrie et idempotence de $\left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{M} \right)$. Nous avons ainsi la factorisation :

$$\mathbf{B} = (\mathbf{V}^b) (\mathbf{V}^b)^T, \quad (3.24a)$$

$$\mathbf{V}^b = \mathbf{E}^b \frac{\mathbf{U}_M}{\sqrt{m}} \in \mathbb{R}^{n(L+1) \times m}, \quad (3.24b)$$

$$m = M - 1. \quad (3.24c)$$

La matrice de déviation \mathbf{V}^b est injective si et seulement si \mathbf{E}^b est injective. Nous supposons dans la suite que c'est le cas. Cette propriété est systématiquement satisfaite par les méthodes d'assimilation de données séquentielles fiables.

Plus généralement, nous avons aussi l'expression compacte suivante de la moyenne empirique \mathbf{v} et de la matrice de déviation \mathbf{V} d'un ensemble \mathbf{E} :

$$\left[\begin{array}{c|c} \mathbf{v} & \mathbf{V} \end{array} \right] = \mathbf{E} \left[\begin{array}{c|c} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} \end{array} \right]. \quad (3.25)$$

Réciproquement, un ensemble \mathbf{E} ayant \mathbf{v} pour moyenne empirique et $\mathbf{V}\mathbf{V}^T$ pour matrice de covariance empirique peut être engendré avec :

$$\mathbf{E} = \left[\begin{array}{c|c} \mathbf{v} & \mathbf{V} \end{array} \right] \left[\begin{array}{c|c} \mathbf{1}_M & \sqrt{m}\mathbf{U}_M \end{array} \right]^T, \quad (3.26)$$

parce que $\left[\begin{array}{c|c} \mathbf{1}_M & \sqrt{m}\mathbf{U}_M \end{array} \right]^T$ est l'inverse matriciel de $\left[\begin{array}{c|c} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} \end{array} \right]$.

Les erreurs d'estimation (ou d'échantillonnage) sont négligées pour construire une

représentation de \mathbf{v} dans le sous-espace de l'ensemble. D'après l'annexe A.4, il existe ainsi une unique variable aléatoire \mathbf{w} à valeurs dans \mathbb{R}^m telle que :

$$\mathbf{v} = \mathbf{v}^b + \mathbf{V}^b \mathbf{w}, \quad (3.27a)$$

$$\mathbb{E}_{p(\mathbf{w}|\mathbf{y}_{:L-S})}[\mathbf{w}] = \mathbf{0}_m, \quad (3.27b)$$

$$\mathbb{C}_{p(\mathbf{w}|\mathbf{y}_{:L-S})}[\mathbf{w}] = \mathbf{I}_m, \quad (3.27c)$$

avec $m = M - 1$ la taille effective de l'ensemble. L'Eq. (3.27a) avec l'Eq. (3.8) forment un changement de variable bien défini entre la trajectoire des états $\mathbf{x}_{0:L}$ et la variable \mathbf{w} :

$$\mathbf{x}_{0:G} = \mathbf{v}_{0:G}^b + \mathbf{V}_{0:G}^b \mathbf{w}, \quad (3.28a)$$

$$\mathbf{x}_l = \mathcal{M}_l(\mathbf{x}_{l-1}) + \mathbf{v}_l^b + \mathbf{V}_l^b \mathbf{w}, \quad (3.28b)$$

pour $l = G + 1, \dots, L$. Les équations 3.28 sont écrites de manière compacte avec :

$$\mathbf{x}_{0:L} = \mathcal{X}_{0:L}(\mathbf{w}). \quad (3.29)$$

3.3.2 Analyse

A l'aide de la règle de Bayes et de l'indépendance conditionnelle comme dans Eq. (3.7a), la densité a posteriori de la variable d'ensemble se factorise en :

$$p(\mathbf{w}|\mathbf{y}_{:L}) \propto p(\mathbf{y}_{K:L}|\mathbf{w}) p(\mathbf{w}|\mathbf{y}_{:L-S}). \quad (3.30)$$

En utilisant l'approximation gaussienne de la densité a priori de moments donnés par Eq. (3.27b,3.27c), on obtient la fonction de coût de la variable d'ensemble :

$$\begin{aligned} J(\mathbf{w}; \mathbf{y}_{K:L}) &\simeq -\ln p(\mathbf{w}|\mathbf{y}_{:L}), \\ J(\mathbf{w}; \mathbf{y}_{K:L}) &\propto \frac{1}{2} \sum_{l=K}^L \|\mathbf{y}_l - \mathcal{H}_l \circ \mathcal{X}_l(\mathbf{w})\|_{\mathbf{R}_l^{-1}}^2 + \frac{1}{2} \|\mathbf{w}\|^2. \end{aligned} \quad (3.31)$$

Les outils utilisés pour minimiser cette fonction de coût sont à présent développés.

3.3.2.1 Analyse linéaire de la variable d'ensemble

Si les opérateurs :

$$\mathcal{F}_l = \mathcal{H}_l \circ \mathcal{X}_l, \quad (3.32)$$

sont affines pour $l = K, \dots, L$; c'est à dire pour tous $\mathbf{w}, \delta\mathbf{w} \in \mathbb{R}^m$:

$$\mathcal{F}_l(\mathbf{w} + \delta\mathbf{w}) = \mathcal{F}_l(\mathbf{w}) + \mathbf{F}_l\delta\mathbf{w}, \quad (3.33)$$

alors la fonction de coût Eq. (3.31) est quadratique. La moyenne et la matrice de covariance a posteriori sont alors algébriquement données par le minimiseur et l'inverse de la matrice hessienne de cette fonction de coût respectivement.

En effet, la matrice hessienne $\mathbf{C}^a = \nabla_{\mathbf{w}}^2 J(\mathbf{w}; \mathbf{y}_{K:L}) > \mathbf{I}_m$ est constante, symétrique, définie-positive. Donc la fonction de coût Eq. (3.31) possède un unique minimiseur :

$$\mathbf{w}^a = \mathbf{w} - (\mathbf{C}^a)^{-1} \nabla_{\mathbf{w}} J(\mathbf{w}; \mathbf{y}_{K:L}). \quad (3.34)$$

D'après un développement de Taylor d'ordre deux en \mathbf{w}^a , cette fonction de coût s'écrit :

$$J(\mathbf{w}; \mathbf{y}_{K:L}) \propto \frac{1}{2} \|\mathbf{w} - \mathbf{w}^a\|_{\mathbf{C}^a}^2. \quad (3.35)$$

La densité gaussienne a posteriori associée est alors :

$$p(\mathbf{w}|\mathbf{y}_{:L}) = \mathcal{N}(\mathbf{w}^a, (\mathbf{C}^a)^{-1}). \quad (3.36)$$

3.3.2.2 Construction récursive des dérivées

Pour effectuer les calculs précédents, les dérivées d'ordre un et deux de la fonction de coût de la variable d'ensemble sont nécessaires. Une construction récursive de ces dérivées est proposée ici.

La fonction de coût de la variable d'ensemble Eq. (3.31) a la structure récursive suivante :

$$J(\mathbf{w}; \mathbf{y}_{K:l}) \propto J(\mathbf{w}; \mathbf{y}_{K:l-1}) + \frac{1}{2} \|\mathbf{y}_l - \mathcal{F}_l(\mathbf{w})\|_{\mathbf{R}_l^{-1}}^2, \quad (3.37)$$

pour $l = K, \dots, L$. Les dérivées première et seconde par rapport à \mathbf{w} de cette expression sont :

$$\nabla_{\mathbf{w}} J(\mathbf{w}; \mathbf{y}_{K:l}) = \nabla_{\mathbf{w}} J(\mathbf{w}; \mathbf{y}_{K:l-1}) - \mathbf{F}_l^T \mathbf{R}_l^{-1} (\mathbf{y}_l - \mathcal{F}_l(\mathbf{w})), \quad (3.38a)$$

$$\nabla_{\mathbf{w}}^2 J(\mathbf{w}; \mathbf{y}_{K:l}) = \nabla_{\mathbf{w}}^2 J(\mathbf{w}; \mathbf{y}_{K:l-1}) + \mathbf{F}_l^T \mathbf{R}_l^{-1} \mathbf{F}_l. \quad (3.38b)$$

Ces équations peuvent être répétées S fois avec la condition initiale :

$$\nabla_{\mathbf{w}} J(\mathbf{w}; \mathbf{y}_{K:l}) = \mathbf{w}, \quad (3.39a)$$

$$\nabla_{\mathbf{w}}^2 J(\mathbf{w}; \mathbf{y}_{K:l}) = \mathbf{I}_m, \quad (3.39b)$$

pour construire ces dérivées de manière récursive.

Nous avons supposé que les opérateurs de cette section étaient affines. La section suivante est dédiée aux méthodes variationnelles lorsque ce caractère linéaire fait défaut.

3.3.2.3 Analyse variationnelle de la variable d'ensemble

L'approche de la méthode de Gauss-Newton pour minimiser les fonctions de coût non quadratiques Björck (1996) nécessite de linéariser les opérateurs. Dans Eq. (3.31), les opérateurs $(\mathcal{F}_l)_{K \leq l \leq L}$ sont linéarisés autour d'un point \mathbf{w}^j obtenu itérativement :

$$\mathcal{F}_l(\mathbf{w}) \simeq \mathbf{f}_l^j + \mathbf{F}_l^j (\mathbf{w} - \mathbf{w}^j), \quad (3.40)$$

où \mathbf{f}_l^j est la valeur en \mathbf{w}^j de l'approximation affine de \mathcal{F}_l et \mathbf{F}_l^j sa partie linéaire pour $l = K, \dots, L$. Avec cette approximation affine, l'assimilation est conduite comme dans la section 3.3.2.1 afin d'obtenir \mathbf{w}^{j+1} et \mathbf{C}^{j+1} . Le processus est itéré depuis l'initialisation $\mathbf{w}^0 = \mathbf{0}_m$ suggérée par la densité a priori jusqu'à ce que le critère de convergence (incomplet) :

$$\|\mathbf{w}^{j+1} - \mathbf{w}^j\| \leq \Delta, \quad (3.41)$$

soit satisfait. Dans ce cas, un point quasiment stationnaire \mathbf{w}^* et une approximation de la hessienne en ce point \mathbf{C}^* sont obtenus. Ce point est une approximation de minimum local strict à la condition suffisante que la hessienne exacte de la fonction de coût en \mathbf{w}^* soit symétrique, définie-positive. Quand bien même, ce minimum peut ne pas être global. La méthode est cependant efficace en pratique pour des problèmes moyennement non linéaires (Dennis and Schnabel, 1996).

La linéarisation Eq. (3.33) peut être vue comme la solution d'un problème d'interpolation linéaire :

$$\mathcal{F}_l^j(\mathbf{D}^j) = \mathbf{f}_l^j \mathbf{1}_M^T + \mathbf{F}_l^j (\mathbf{D}^j - \mathbf{w}^j \mathbf{1}_M^T), \quad (3.42)$$

sur un ensemble \mathbf{D}^j de points d'interpolation. Si cet ensemble est engendré depuis une moyenne empirique \mathbf{w}^j et une matrice de déviation \mathbf{W}^j comme dans Eq. (3.26) :

$$\mathbf{D}^j = \begin{bmatrix} \mathbf{w}^j & \mathbf{W}^j \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T, \quad (3.43)$$

la solution du problème d'interpolation linéaire est alors :

$$\begin{bmatrix} \mathbf{f}_l^j & \mathbf{F}_l^j \end{bmatrix} = \mathcal{F}_l^j(\mathbf{D}^j) \begin{bmatrix} \frac{1}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} (\mathbf{W}^j)^{-1} \end{bmatrix}. \quad (3.44)$$

Dans les deux paragraphes suivants, des définitions de \mathbf{W}^j Eq. (3.43) inspirées de Boc-

quet and Sakov (2012); Sakov et al. (2012) sont proposées.

Méthode bundle Si $\mathbf{W}^j = \epsilon \mathbf{I}_m$, alors Eq. (3.44) représente les différences finies de \mathcal{F}_l de pas ϵ :

$$\begin{bmatrix} \mathbf{f}_l^j & \mathbf{F}_l^j \end{bmatrix} \simeq \begin{bmatrix} \mathcal{F}_l(\mathbf{w}^j) & \frac{d\mathcal{F}_l}{d\mathbf{w}}(\mathbf{w}^j) \end{bmatrix}. \quad (3.45)$$

La linéarisation Eq. (3.40) est un développement de Taylor d'ordre 1 de \mathcal{F}_l .

Méthode transform Si \mathbf{W}^j est tel que :

$$(\mathbf{W}^j) (\mathbf{W}^j)^\top = (\mathbf{C}^j)^{-1}, \quad (3.46)$$

alors Eq. (3.44) est, d'après l'annexe A.3.1, la meilleure approximation affine de \mathcal{F}_l au sens de l'erreur quadratique moyenne suivant une variable aléatoire $\tilde{\mathbf{w}}$ ayant \mathbf{w}^j pour moyenne et $(\mathbf{W}^j) (\mathbf{W}^j)^\top$ pour matrice de covariance.

Pour évaluer $\mathcal{F}_l(\mathbf{D}^j)$ efficacement, la structure récursive de $\mathcal{X}_{0:L}$ Eq. (3.28) peut être exploitée dans le schéma d'évaluation suivant :

$$\mathbf{E}_{0:G}^j = \mathbf{v}_{0:G}^b \mathbf{1}_M^\top + \mathbf{V}_{0:G} \mathbf{D}^j, \quad (3.47a)$$

$$\mathbf{E}_l^j = \mathcal{M}_l(\mathbf{E}_{l-1}^j) + \mathbf{v}_l^b \mathbf{1}_M^\top + \mathbf{V}_l \mathbf{D}^j, \quad (3.47b)$$

$$\mathcal{F}_l(\mathbf{D}^j) = \mathcal{H}_l(\mathbf{E}_l^j), \quad (3.47c)$$

pour $l = G + 1, \dots, L$. Il requiert un total de $(L - G) \times M$ évaluations du modèle à chaque itération j . Notons que, puisque $\mathbf{E}_{0:L}^j = \mathcal{X}_{0:L}(\mathbf{D}^j)$ est aussi évalué dans ce schéma; l'application de 3.44 à $\mathcal{X}_{0:L}$ en produit une linéarisation.

Dans la section suivante, le nombre d'évaluations du modèle, qui est critique pour des applications géophysiques de haute dimension, est davantage réduit.

3.3.2.4 Décomposition de la matrice de déviation et modèle de substitution

Les évaluations du modèle dans des directions associées avec une petite variance ne requiert que peu de précision. Dans cette section, les directions de \mathbf{E}_l^j Eq. (3.47) sont divisées en deux paquets. Le modèle est évalué sur le premier et un modèle de substitution moins coûteux est évalué sur le second. Cette séparation est particulièrement efficace avec une dynamique instable car les directions de l'espace des états sont associées à des variances relativement différentes.

Soient $\begin{bmatrix} \mathbf{x}_l^j & \mathbf{X}_l^j \end{bmatrix}$ la moyenne empirique et une matrice de déviation de \mathbf{E}_l^j calculé

avec Eq. (3.47) :

$$\begin{bmatrix} \mathbf{x}_l^j & \mathbf{X}_l^j \end{bmatrix} = \mathbf{E}_l^j \begin{bmatrix} \frac{\mathbf{1}_M}{M} & \frac{\mathbf{U}_M}{\sqrt{m}} \end{bmatrix}. \quad (3.48)$$

Une décomposition de la matrice de déviation est effectuée :

$$\mathbf{X}_l^j = \begin{bmatrix} \mathbf{X}_l^{0,j} & \mathbf{X}_l^{1,j} \end{bmatrix} \begin{bmatrix} \mathbf{V}^0 & \mathbf{V}^1 \end{bmatrix}^T, \quad (3.49)$$

où la matrice $\begin{bmatrix} \mathbf{V}^0 & \mathbf{V}^1 \end{bmatrix}$ est orthogonale et $\mathbf{X}_l^{s,j} \in \mathbb{R}^{n \times m^s}$, $\mathbf{V}^s \in \mathbb{R}^{m \times m^s}$ pour $s = 0, 1$ avec $m^0 + m^1 = m$. Une telle décomposition peut provenir d'une décomposition en valeurs singulières (SVD, e.g. [Stoer and Bulirsch, 1982](#)) ou exploiter d'éventuelles colonnes nulles de \mathbf{X}_l^j . Si \mathcal{M}_{l+1} est affine de partie linéaire \mathbf{M}_{l+1} la décomposition précédente Eq. (3.49) mène à :

$$\begin{aligned} \mathcal{M}_{l+1}(\mathbf{E}_l^j) &= \mathcal{M}_{l+1}(\mathbf{x}_l^j) \mathbf{1}_M^T \\ &+ \left(\mathbf{M}_{l+1} \mathbf{X}_l^{0,j} \mathbf{V}^{0T} + \mathbf{M}_{l+1} \mathbf{X}_l^{1,j} \mathbf{V}^{1T} \right) \sqrt{m} \mathbf{U}_M^T, \end{aligned} \quad (3.50)$$

où l'on a appliqué Eq. (3.26) à \mathbf{E}_l^j . Dans l'expression Eq. (3.50), $\mathcal{M}_{l+1}(\mathbf{x}_l^j)$ et $\mathbf{M}_{l+1} \mathbf{X}_l^{0,j}$ peuvent être vus comme la moyenne empirique et la matrice de déviation de la propagation d'un ensemble $\mathcal{M}_{l+1}(\mathbf{E}_l^{0,j})$; alors que $\mathbf{M}_{l+1} \mathbf{X}_l^{1,j}$ peut être vu comme la matrice de déviation de la propagation d'un autre ensemble $\mathcal{M}_{l+1}(\mathbf{E}_l^{1,j})$. Les ensembles $\mathbf{E}_l^{0,j}$, $\mathbf{E}_l^{1,j}$ sont tels que :

$$\mathbf{E}_l^{s,j} = \begin{bmatrix} \mathbf{x}_l^j & \mathbf{X}_l^{s,j} \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M^s} & \sqrt{m^s} \mathbf{U}_{M^s} \end{bmatrix}^T \in \mathbb{R}^{n \times M^s}, \quad (3.51)$$

avec $M^s = m^s + 1$ pour $s = 0, 1$.

Lorsque le modèle \mathcal{M}_{l+1} et son substitut $\tilde{\mathcal{M}}_{l+1}$ sont respectivement évalués sur $\mathbf{E}_l^{0,j}$ et $\mathbf{E}_l^{1,j}$, le prochain ensemble \mathbf{E}_{l+1}^j Eq. (3.47b) peut être recomposé en imitant Eq. (3.50) avec :

$$\begin{aligned} \mathbf{E}_{l+1}^j &\simeq \begin{bmatrix} \mathcal{M}_{l+1}(\mathbf{E}_l^{0,j}) & \tilde{\mathcal{M}}_{l+1}(\mathbf{E}_l^{1,j}) \end{bmatrix} \begin{bmatrix} \frac{\mathbf{1}_{M^0}}{M^0} & \frac{\mathbf{U}_{M^0}}{\sqrt{m^0}} \mathbf{V}^{0T} \\ \mathbf{0}_{M^1} & \frac{\mathbf{U}_{M^1}}{\sqrt{m^1}} \mathbf{V}^{1T} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T \\ &+ \mathbf{v}_{l+1}^b \mathbf{1}_M^T + \mathbf{V}_{l+1}^b \mathbf{D}_{l+1}^j. \end{aligned} \quad (3.52)$$

Cette procédure d'évaluation requiert un total de $(L - G) \times M^0$ évaluations du modèle et $(L - G) \times M^1$ évaluations de son substitut à chaque itération j . La méthode est aussi applicable à l'opérateur d'observation et est généralisable à plus d'un modèle de substitution.

3.3.3 Propagation

L'objectif de cette section est de calculer les statistiques a priori de la variable d'ébauche du prochain cycle $\mathbf{v}'_{S:L+S} = \text{vec}(\mathbf{x}_{S:G+S}, \boldsymbol{\varepsilon}_{G+S+1:L}^q)$ à partir des statistiques a posteriori de $\mathbf{v}_{0:L} = \text{vec}(\mathbf{x}_{0:G}, \boldsymbol{\varepsilon}_{G+1:L}^q)$.

L'analyse variationnelle de la variable d'ensemble produit l'approximation gaussienne a posteriori suivante :

$$p(\mathbf{w}|\mathbf{y}_{:L}) \simeq \mathcal{N}(\mathbf{w}^a, \mathbf{W}^a \mathbf{W}^{aT}), \quad (3.53)$$

où $\mathbf{w}^a = \mathbf{w}^*$ et \mathbf{W}^a est telle que :

$$\mathbf{W}^a \mathbf{W}^{aT} = (\mathbf{C}^*)^{-1}. \quad (3.54)$$

À cause de la relation affine entre \mathbf{w} et $\mathbf{v}_{0:L}$ dans Eq. (3.27a), sa moyenne a posteriori et une de ses matrices de déviation sont :

$$\begin{bmatrix} \mathbf{v}_{0:L}^a & \mathbf{V}_{0:L}^a \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{0:L}^b + \mathbf{V}_{0:L}^b \mathbf{w}^a & \mathbf{V}_{0:L}^b \mathbf{W}^a \end{bmatrix}. \quad (3.55)$$

Grâce à l'indépendance des erreurs du modèle, leurs statistiques peuvent être utilisées pour étendre les statistiques a posteriori de $\mathbf{v}_{0:L}$ en statistiques a posteriori de $\mathbf{v}_{0:L+S} = \text{vec}(\mathbf{v}_{0:L}, \boldsymbol{\varepsilon}_{L+1:L+S}^q)$:

$$\mathbf{v}_{0:L+S}^a = \text{vec}(\mathbf{v}_{0:L}^a, \mathbf{0}_n, \dots, \mathbf{0}_n) \in \mathbb{R}^{n(L+S+1)}, \quad (3.56a)$$

$$\mathbf{V}_{0:L+S}^a = \text{diag}(\mathbf{V}_{0:L}^a, \mathbf{V}_{L+1}^q, \dots, \mathbf{V}_{L+S}^q) \in \mathbb{R}^{n(L+S+1) \times m'}, \quad (3.56b)$$

où $m' = m + \sum_{l=L+1}^{L+S} m_l^q$. Les moyennes $\mathbf{v}_l^q = \mathbf{0}_n$ et matrices de déviations \mathbf{V}_l^q sont ici calculées avec Eq. (3.25) depuis des ensembles \mathbf{E}_l^q échantillonnant l'erreur modèle $\boldsymbol{\varepsilon}_l^q$:

$$\begin{bmatrix} \mathbf{v}_l^q & \mathbf{V}_l^q \end{bmatrix} = \mathbf{E}_l^q \begin{bmatrix} \frac{\mathbf{1}_{M_l^q}}{M_l^q} & \frac{\mathbf{U}_{M_l^q}}{\sqrt{m_l^q}} \end{bmatrix}. \quad (3.57)$$

Les ensembles d'erreurs modèle sont conçus pour avoir peu de membres $M_l^q = m_l^q + 1$ ($< n$). Le réalisme et les avantages de la représentation des covariances d'erreurs modèle $\mathbf{Q}_l \simeq (\mathbf{V}_l^q)(\mathbf{V}_l^q)^T$ avec un nombre limité de modes a été discuté dans la section 4 de Sakov and Bocquet (2018). Les statistiques étendues Eq. (3.56) sont ensuite propagées

de S pas de temps avec le schéma d'évaluation suivant :

$$\begin{aligned} \mathbf{E}'_{0:G}{}^b &= \begin{bmatrix} \mathbf{v}_{0:G}^a & \mathbf{V}_{0:G}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M'} & \sqrt{m'}\mathbf{U}_{M'} \end{bmatrix}^T, \\ \mathbf{E}'_{l+1}{}^b &= \mathcal{M}_{l+1}(\mathbf{E}'_l{}^b) + \begin{bmatrix} \mathbf{v}_l^a & \mathbf{V}_l^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M'} & \sqrt{m'}\mathbf{U}_{M'} \end{bmatrix}^T, \\ \mathbf{E}'_{G+S+1:L+S}{}^b &= \begin{bmatrix} \mathbf{v}_{G+S+1:L+S}^a & \mathbf{V}_{G+S+1:L+S}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M'} & \sqrt{m'}\mathbf{U}_{M'} \end{bmatrix}^T, \end{aligned}$$

pour $l = G, \dots, G + S - 1$. Notons que le nombre d'évaluations du modèle peut être réduit avec les décompositions de matrices de déviation et les modèles de substitution de la section 3.3.2.4. En particulier, avec Eq. (3.55) et la définition de $\mathcal{X}_{0:L}$ Eq. (3.28) on a :

$$\mathbf{E}'_{0:L}{}^b = \mathcal{X}_{0:L}(\mathbf{D}'^a), \quad (3.59a)$$

$$\mathbf{D}'^a = \begin{bmatrix} \mathbf{w}^a & \begin{bmatrix} \mathbf{W}^a & \mathbf{0}_{m \times (m'-m)} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M'} & \sqrt{m'}\mathbf{U}_{M'} \end{bmatrix}^T. \quad (3.59b)$$

La linéarisation de $\mathcal{X}_{0:L}$ provenant de l'analyse peut ainsi être utilisée pour réduire le coût de calcul de $\mathbf{E}'_{0:L}{}^b$.

Après cette propagation, l'ensemble $\mathbf{E}'_{S:L+S}{}^b$ échantillonne $\mathbf{v}'_{S:L+S} = \text{vec}(\mathbf{x}_{S:G+S}, \boldsymbol{\varepsilon}_{G+S+1:L+S}^a)$ connaissant $\mathbf{y}_{:L}$. Malheureusement, $\mathbf{E}'_{S:L+S}{}^b$ possède $M' = m' + 1$ membres au lieu de $M = m + 1$ comme $\mathbf{E}'_{0:L}$ au début du cycle. Dans la section suivante, cette dimension est réduite.

3.3.3.1 Réduction de la dimension et marginalisation

La matrice de déviation de la variable d'ébauche au prochain cycle $\mathbf{V}'_{S:L+S}$ doit engendrer la même matrice de covariance que la matrice de covariance empirique de $\mathbf{E}'_{S:L+S}{}^b$:

$$\left(\mathbf{V}'_{S:L+S}\right) \left(\mathbf{V}'_{S:L+S}\right)^T \simeq \mathbf{B}'_{S:L+S}, \quad (3.60)$$

en ayant seulement m colonnes. Cette équation est approchée parce qu'une solution exacte $\mathbf{V}'_{S:L+S}$ dans $\mathbb{R}^{nL \times m}$ n'existe pas toujours. Dans ce cas, une réduction de dimension basée sur une diagonalisation fournit la meilleure approximation en norme de Frobenius (théorème d'Eckart-Young). Cette solution n'est pas unique puisque toute multiplication à droite par une matrice orthogonale constitue une nouvelle solution. Choisir cette matrice orthogonale aléatoirement (e.g., par décomposition QR d'une matrice gaussienne) peut améliorer la précision (Sakov et al., 2018).

Cependant, le schéma d'évaluation Eq. (3.58) ne modifie pas la structure bloc diagonale Eq. (3.56b) après l'indice $G + S$. Ainsi, la matrice de covariance empirique de

$\mathbf{E}'_{S:L+S}{}^b$ a la structure bloc diagonale :

$$\mathbf{B}'_{S:L+S} = \text{diag} \left(\mathbf{B}'_{S:D+S}, \mathbf{V}_{D+S+1}^q \mathbf{V}_{D+S+1}^{qT}, \dots, \mathbf{V}_{L+S}^q \mathbf{V}_{L+S}^{qT} \right), \quad (3.61)$$

avec $D = \max(G, K - 1)$. Une alternative plus efficace est alors de procéder à la réduction de dimension sur le premier bloc diagonal seul :

$$\left(\mathbf{V}'_{S:D+S}{}^b \right) \left(\mathbf{V}'_{S:D+S}{}^b \right)^T \simeq \mathbf{B}'_{S:D+S}, \quad (3.62)$$

avec $\mathbf{V}'_{S:D+S}{}^b \in \mathbb{R}^{n(D+1) \times m^b}$ et $m^b = m - \sum_{l=D+S+1}^{L+S} m_l^q$. Finalement, la matrice de déviation résultante est étendue à l'aide des matrices de déviation des erreurs modèle :

$$\mathbf{V}'_{S:L+S}{}^b = \text{diag} \left(\mathbf{V}'_{S:L+S}{}^b, \mathbf{V}_{D+S+1}^q, \dots, \mathbf{V}_{L+S}^q \right). \quad (3.63)$$

En plus d'un avantage en termes de calculs, cette réduction de la dimension ne compare pas les valeurs propres de $\mathbf{B}'_{S:D+S}$ avec les valeurs propres des $(\mathbf{V}_l^q)_{D+S+1 \leq l \leq L+S}$. Puisque les premières sont analysées, elles sont probablement plus petites et donc abandonnées si la diagonalisation est faite sur la matrice de covariance complète $\mathbf{B}'_{S:L+S}$.

Dans la section suivante, le pseudo-code de l'IEnKS-Q et plusieurs variantes sont discutés.

3.4 Algorithmes

Dans cette section, une implémentation de l'IEnKS-Q basée sur la construction de la section 3.3 est proposée. L'efficacité numérique est ensuite considérée dans plusieurs variantes.

L'algorithme 3.1 est le pseudo-code d'un cycle de la variante la plus simple de l'IEnKS-Q. Elle est basée sur une représentation 4D de la plupart des objets mathématiques de la fenêtre d'assimilation. Il peut être vu comme une application directe de l'IEnKS sur une variable d'état étendue temporellement.

Le cycle démarre avec l'analyse de la section 3.3.2. La ligne 1 est la moyenne et la matrice de déviation initiale de la variable d'ensemble. Cette matrice de déviation est contractée par un facteur ϵ , le pas des différences finies, dans la méthode bundle. La boucle **repeat** lignes 2-16 correspond aux itérations de Gauss-Newton de la section 3.3.2.3. La ligne 3 initialise le gradient de la fonction de coût et la matrice hessienne approchée comme dans Eq. (3.39). La ligne 4 initialise l'ensemble 4D avec les statistiques de la variable d'ébauche et de la variable d'ensemble pour évaluer le schéma Eq. (3.47).

La boucle **for** lignes 5-13 construit les dérivées de la fonction de coût récursivement comme dans la section 3.3.2.2. Quand $l > G$, la ligne 7 incrémente l'ensemble à t_l avec la propagation de l'ensemble à t_{l-1} . Quand $l \geq K$, la ligne 10 linéarise l'opérateur \mathcal{F}_l comme dans Eq. (3.44) et la ligne 11 incrémente les dérivées de la fonction de coût suivant Eq. (3.44). La ligne 14 résout l'équation de gradient nulle Eq. (3.34) et incrémente le minimiseur. La ligne 15 met à jour la matrice de déviation de la variable d'ensemble dans le cas de la version transform comme dans Eq. (3.46) à l'aide d'une factorisation de Cholesky (e.g., Stoer and Bulirsch, 1982) de l'inverse de l'approximation de la matrice hessienne. La ligne 16 est le critère de convergence de Gauss-Newton Eq. (3.41). Puis l'étape de propagation section 3.7b commence. La ligne 17 met à jour la matrice de déviation de la variable d'ensemble Eq. (3.54) avec une factorisation de Cholesky de l'inverse de l'approximation de la matrice hessienne. La ligne 18 calcule les statistiques a posteriori de la variable d'ébauche et les étend avec les statistiques d'erreur modèle comme dans Eq. (3.55,3.56). La ligne 19 engendre l'ensemble étendu associé à ces statistiques avec Eq. (3.26). La boucle **for** lignes 20-22 propage les statistiques a posteriori de l'état de S pas de temps depuis G à l'aide du schéma Eq. (3.58). La ligne 21 incrémente l'ensemble à t_l avec la propagation de l'ensemble à t_{l-1} . La ligne 23 calcule la moyenne et la matrice de déviation de la variable d'ébauche du prochain cycle avec Eq. (3.25). La ligne 24 est la réduction de dimension de la section 3.3.3.1. La ligne 25 dilate la matrice de déviation.

Comme discuté dans le dernier paragraphe de la section 3.3.3.1, la réduction de la dimension peut être améliorée en la restreignant au premier bloc diagonal de la matrice de covariance $\mathbf{B}'_{S:L+S}$ Eq. (3.61). L'algorithme 3.2 remplace la réduction de la dimension de l'algorithme 3.1 conformément.

Comme discuté Eq. (3.59), les linéarisations du modèle provenant de l'analyse peuvent être utilisées pour réduire le nombre d'évaluations du modèle lors de la propagation. Ceci est implémenté dans l'algorithme 3.3.

La résolution du système ligne 14 et les factorisations matricielles 15, 17 peuvent bénéficier d'une factorisation de Cholesky. Si les colonnes de \mathbf{C} sont permutées, les inversions matricielles peuvent s'effectuer sur les facteurs de Cholesky qui sont triangulaires inférieurs. Ceci est plus efficace que d'inverser \mathbf{C} puis de calculer ses facteurs de Cholesky. Le détail des opérations est donné dans l'algorithme 3.4.

L'algorithme 3.4 nécessite un nombre d'évaluations du modèle qui est proportionnel au carré de la longueur de la fenêtre d'assimilation. Ce nombre peut être divisé de moitié en exploitant les colonnes nulles de la matrice de déviation de \mathbf{E}_{l-1} . Si $D = \max(G, K - 1)$ alors $\mathbf{V}_{0:L}^b = \text{diag}(\mathbf{V}_{0:D}^b, \mathbf{V}_{D+1}^q, \dots, \mathbf{V}_L^q)$ est bloc diagonale. De plus, si

\mathbf{W} est triangulaire inférieure alors la matrice de déviation $\mathbf{X}_{0:L} = \mathbf{V}_{0:L}^b \mathbf{W}$ de $\mathbf{E}_{0:L}$ de la ligne 4 de l'algorithme 3.1 est aussi bloc diagonale. Ainsi ses lignes vérifient la décomposition $\mathbf{X}_{l-1} = \begin{bmatrix} \mathbf{X}_{l-1}^0 & \mathbf{0}_{n \times (m-m^0)} \end{bmatrix}$ avec $\mathbf{X}_{l-1}^0 \in \mathbb{R}^{n \times m^0}$ et $m^0 = m' - \sum_{k=\max(D,l-1)+1}^{L+S} m_l^q$. Le schéma d'évaluation de l'algorithme 3.5 utilise cette décomposition de la matrice de déviation avec Eq. (3.52) pour réduire le nombre d'évaluations du modèle.

L'IEEnKS-Q avec la variante de l'algorithme 3.5 requiert encore un nombre d'évaluations du modèle proportionnel au carré de la longueur de la fenêtre d'assimilation. Des décompositions de matrice de déviation basées sur des SVD limitées et des modèles de substitution peuvent aider à réduire d'avantage le nombre d'évaluations du modèle. Le schéma d'évaluation général basé sur les résultats de la section 3.3.2.4 est donné dans l'algorithme 3.6

Algorithm 3.1 Un cycle de l'EnKS-Q bundle/*transform***Require:** Les paramètres de fenêtre $0 \leq G \leq K \leq L$ et $S = L - K + 1$.Les observations $\mathbf{y}_{K:L}$.Le modèle \mathcal{M} et l'opérateur d'observation \mathcal{H} .La moyenne et matrice de déviation a priori $\mathbf{v}_{0:L}^b \in \mathbb{R}^{n \times (L+1)}$, $\mathbf{V}_{0:L}^b \in \mathbb{R}^{n \times (L+1) \times m}$.Les matrices de déviation d'erreur modèle $\mathbf{V}_l^q \in \mathbb{R}^{n \times m_l^q}$ for $l = L + 1, \dots, L + S$.La dimension étendue $m' = m + \sum_{l=L+1}^{L+S} m_l^q$, $M' = m' + 1$.Des matrices orthogonales $\begin{bmatrix} \frac{1_N}{\sqrt{N}} & \mathbf{U}_N \end{bmatrix} \in \mathbb{R}^{N \times N}$ for $N = M, M'$.Le pas des différences finies $\epsilon = 1$, le critère d'arrêt de Gauss-Newton Δ , l'inflation λ .**Ensure:** $\mathbf{v}_{S:L+S}^b \in \mathbb{R}^{n \times (L+1)}$, $\mathbf{V}_{S:L+S}^b \in \mathbb{R}^{n \times (L+1) \times m}$ sont la moyenne et la matrice de déviation a priori du prochain cycle.—**Analysis**—

- 1: $\begin{bmatrix} \mathbf{w} & \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_m & \epsilon \mathbf{I}_m \end{bmatrix}$
- 2: **repeat**
- 3: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} = \begin{bmatrix} \mathbf{w} & \mathbf{I}_m \end{bmatrix}$
- 4: $\mathbf{E}_{0:L} = \begin{bmatrix} \mathbf{v}_{0:L}^b + \mathbf{V}_{0:L}^b \mathbf{w} & \mathbf{V}_{0:L}^b \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{1}_M & \sqrt{m} \mathbf{U}_M \end{bmatrix}^T$
- 5: **for** $l = 0 \dots L$ **do**
- 6: **if** $l > G$ **then**
- 7: $\mathbf{E}_l += \mathcal{M}_l(\mathbf{E}_{l-1})$
- 8: **end if**
- 9: **if** $l \geq K$ **then**
- 10: $\begin{bmatrix} \mathbf{f} & \mathbf{F} \end{bmatrix} = \mathcal{H}_l(\mathbf{E}_l) \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \mathbf{W}^{-1} \end{bmatrix}$
- 11: $\begin{bmatrix} \mathbf{g} & \mathbf{C} \end{bmatrix} += \mathbf{F}^T \mathbf{R}_l^{-1} \begin{bmatrix} \mathbf{f} - \mathbf{y}_l & \mathbf{F} \end{bmatrix}$
- 12: **end if**
- 13: **end for**
- 14: solve $\mathbf{C} \times \delta \mathbf{w} = -\mathbf{g}$ for $\delta \mathbf{w} \in \mathbb{R}^m$, $\mathbf{w} += \delta \mathbf{w}$
- 15: solve $\mathbf{W} \mathbf{W}^T = \mathbf{C}^{-1}$ for $\mathbf{W} \in \mathbb{R}^{m \times m}$ lower triangular
- 16: **until** $\|\delta \mathbf{w}\| \leq \Delta$

—**Propagation**—

- 17: solve $\mathbf{W}^a \mathbf{W}^{aT} = \mathbf{C}^{-1}$ for $\mathbf{W}^a \in \mathbb{R}^{m \times m}$ lower triangular
- 18: $\begin{bmatrix} \mathbf{v}_{0:L+S}^a & \mathbf{V}_{0:L+S}^a \end{bmatrix} = \begin{bmatrix} \text{vec}(\mathbf{v}_{0:L}^b + \mathbf{V}_{0:L}^b \mathbf{w}, \mathbf{0}_{n \times S}), & \text{diag}(\mathbf{V}_{0:L}^b \mathbf{W}^a, \mathbf{V}_{L+1}^q, \dots, \mathbf{V}_{L+S}^q) \end{bmatrix}$
- 19: $\mathbf{E}_{0:L+S}^a = \begin{bmatrix} \mathbf{v}_{0:L+S}^a & \mathbf{V}_{0:L+S}^a \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M'} & \sqrt{m'} \mathbf{U}_{M'} \end{bmatrix}^T$
- 20: **for** $l = G + 1, \dots, G + S$ **do**
- 21: $\mathbf{E}_l^a += \mathcal{M}_l(\mathbf{E}_{l-1}^a)$
- 22: **end for**
- 23: $\begin{bmatrix} \mathbf{v}_{S:L+S}^b & \mathbf{V}_{S:L+S}^b \end{bmatrix} = \mathbf{E}_{S:L+S}^a \begin{bmatrix} \frac{1}{M'} \mathbf{1}_{M'} & \frac{1}{\sqrt{m'}} \mathbf{U}_{M'} \end{bmatrix}$
- 24: solve $(\mathbf{V}_{S:L+S}^b) (\mathbf{V}_{S:L+S}^b)^T \simeq (\mathbf{V}_{S:L+S}^b) (\mathbf{V}_{S:L+S}^b)^T$ for $\mathbf{V}_{S:L+S}^b \in \mathbb{R}^{n \times K \times m}$
- 25: $\mathbf{V}_{S:L+S}^b = \lambda \mathbf{V}_{S:L+S}^b$

Algorithm 3.2 Réduction de la dimension par bloc

Require: $D = \max(G, K - 1)$.La dimension réduite $m^b = m - \sum_{l=D+S+1}^{L+S} m_l^q$.**Ensure:** Remplace la réduction de dimension lignes 24 de l'algorithme 3.1.

- 1: solve $(\mathbf{V}_{S:D+S}^b)(\mathbf{V}_{S:D+S}^b)^T \simeq (\mathbf{V}_{S:D+S}^b)(\mathbf{V}_{S:D+S}^b)^T$ for $\mathbf{V}_{S:D+S}^b \in \mathbb{R}^{n \times K \times m^b}$
with an eigendecomposition
 - 2: $\mathbf{V}_{S:L+S}^b = \text{diag}(\mathbf{V}_{S:D+S}^b, \mathbf{V}_{D+S+1}^q, \dots, \mathbf{V}_{L+S}^q)$
-

Algorithm 3.3 Propagation linéarisée

Ensure: Remplace la ligne 21 de l'algorithme 3.1.

- 1: **if** $l \leq L$ **then**
 - 2: $\mathbf{E}_l^a = \mathbf{E}_l \left(\frac{1}{M} \mathbf{1}_M \mathbf{1}_{M'}^T + \sqrt{\frac{m'}{m}} \mathbf{U}_M \mathbf{W}^{-1} \mathbf{W}^a \begin{bmatrix} \mathbf{I}_m & \mathbf{0}_{m \times (m'-m)} \end{bmatrix} \mathbf{U}_{M'}^T \right)$
 - 3: **else**
 - 4: $\mathbf{E}_l^a += \mathcal{M}_l(\mathbf{E}_{l-1}^a)$
 - 5: **end if**
-

Algorithm 3.4 Résolution efficace du système linéaire et de la factorisation de l'inverse matricielle avec une décomposition de Cholesky.

Require: La matrice de permutation $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.**Ensure:** Résout les problèmes lignes 14, 15 et lignes 17 dans l'algorithme 3.1

- 1: solve $\mathbf{L}\mathbf{L}^T = \mathbf{P}\mathbf{C}\mathbf{P}$ for $\mathbf{L} \in \mathbb{R}^{m \times m}$ lower triangular with a Cholesky decomposition
 - 2: $\mathbf{W}^{-1} = \mathbf{P}\mathbf{L}^T\mathbf{P}$ is lower triangular
 - 3: invert the triangular matrix \mathbf{L}^T to compute $\mathbf{W} = \mathbf{P}\mathbf{L}^{-T}\mathbf{P}$ if necessary
 - 4: solve the triangular linear system $\mathbf{L}\mathbf{u} = -\mathbf{P}\mathbf{g}$ for $\mathbf{u} \in \mathbb{R}^m$
 - 5: solve the triangular linear system $\mathbf{L}^T\mathbf{v} = \mathbf{u}$ for $\mathbf{v} \in \mathbb{R}^m$
 - 6: $\delta\mathbf{w} = \mathbf{P}\mathbf{v}$
-

Algorithm 3.5 Évaluation du modèle avec troncation des matrices de déviation**Require:** $D = \max(G, K - 1)$.Le nombre de colonnes non nulles $m^0 = m' - \sum_{k=\max(D, l-1)+1}^{L+S} m_k^q$.Une matrice orthogonale $\begin{bmatrix} \frac{\mathbf{1}_{M^0}}{\sqrt{M^0}} & \mathbf{U}_{M^0} \end{bmatrix}$, $M^0 = m^0 + 1$.**Ensure:** Remplace les évaluations du modèle sur les ensembles lignes 7, 21 de l'algorithme 3.1.

- 1: $\begin{bmatrix} \mathbf{x}_{l-1} & \mathbf{X}_{l-1} \end{bmatrix} = \mathbf{E}_{l-1} \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \end{bmatrix}$
- 2: $\mathbf{V}^0 = \begin{bmatrix} \mathbf{I}_{m^0} & \mathbf{0}_{m^0 \times (m-m^0)} \end{bmatrix}^T$
- 3: $\mathbf{E}^0 = \begin{bmatrix} \mathbf{x}_{l-1} & \mathbf{X}_{l-1} \mathbf{V}^0 \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M^0} & \sqrt{m^0} \mathbf{U}_{M^0} \end{bmatrix}^T$
- 4: $\mathbf{E}_l += \mathcal{M}_l(\mathbf{E}^0) \left(\frac{1}{M^0} \mathbf{1}_{M^0} \mathbf{1}_M^T + \sqrt{\frac{m}{m^0}} \mathbf{U}_{M^0} \mathbf{V}^{0T} \mathbf{U}_M^T \right)$

Algorithm 3.6 Évaluation du modèle et de son substitut avec décomposition des matrices de déviation**Require:** Le modèle \mathcal{M}_l et son substitut $\widetilde{\mathcal{M}}_l$.Les ensembles substitut sizes m^s , $M^s = m^s + 1$ for $s = 0, 1$.Des matrices orthogonales $\begin{bmatrix} \frac{\mathbf{1}_N}{\sqrt{N}} & \mathbf{U}_N \end{bmatrix}$ for $N = M, M^0, M^1$.**Ensure:** Remplace les évaluations du modèle lignes 7, 21 de l'algorithme 3.1.

- 1: $\begin{bmatrix} \mathbf{x}_{l-1} & \mathbf{X}_{l-1} \end{bmatrix} = \mathbf{E}_{l-1} \begin{bmatrix} \frac{1}{M} \mathbf{1}_M & \frac{1}{\sqrt{m}} \mathbf{U}_M \end{bmatrix}$
- 2: solve $\begin{bmatrix} \mathbf{X}^0 & \mathbf{X}^1 \end{bmatrix} \begin{bmatrix} \mathbf{V}^0 & \mathbf{V}^1 \end{bmatrix}^T = \mathbf{X}_{l-1}$
for $\begin{bmatrix} \mathbf{X}^0 & \mathbf{X}^1 \end{bmatrix}$ in $\mathbb{R}^{n \times m^0} \times \mathbb{R}^{n \times m^1}$, $\begin{bmatrix} \mathbf{V}^0 & \mathbf{V}^1 \end{bmatrix}$ orthogonal in $\mathbb{R}^{m \times m^0} \times \mathbb{R}^{m \times m^1}$
- 3: **for** $s = 0, 1$ **do**
- 4: $\mathbf{E}^s = \begin{bmatrix} \mathbf{x}_{l-1} & \mathbf{X}^s \end{bmatrix} \begin{bmatrix} \mathbf{1}_{M^s} & \sqrt{m^s} \mathbf{U}_{M^s} \end{bmatrix}^T$
- 5: **end for**
- 6: $\mathbf{E}_l += \begin{bmatrix} \mathcal{M}_l(\mathbf{E}^0) & \widetilde{\mathcal{M}}_l(\mathbf{E}^1) \end{bmatrix} \begin{bmatrix} \frac{1}{M^0} \mathbf{1}_{M^0} \mathbf{1}_M^T + \sqrt{\frac{m}{m^0}} \mathbf{U}_M \mathbf{V}^0 \mathbf{U}_{M^0}^T & \sqrt{\frac{m}{m^1}} \mathbf{U}_M \mathbf{V}^1 \mathbf{U}_{M^1}^T \end{bmatrix}^T$

3.5 Résultats numériques

Dans ce qui suit, des expériences numériques sont effectuées avec le modèle de Lorenz à 40 variables (L96, Lorenz and Emanuel, 1998). Il représente un cercle d'atmosphère de moyenne latitude. Il est décrit par en ensemble de n équations différentielles non linéaires :

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2}) x_{j-1} - x_j + F. \quad (3.64)$$

où x_j est la j ème composante modulo n de \mathbf{x} , $n = 40$, $F = 8$. Cette équation est intégrée à l'aide d'un schéma de Runge-Kutta d'ordre 4 et un pas de temps d'intégration de $\delta t = 0.05$. La dynamique de L96 est chaotique ; son plus grand exposant de Lyapunov est $\simeq 1,7$. Le bruit d'erreur modèle est stochastiquement ajouté à la trajectoire de l'état (vérité) à chaque intégration de Δt du modèle :

$$\mathbf{x}_{l+1} = \mathcal{M}(\mathbf{x}_l) + \boldsymbol{\varepsilon}_{l+1}^q,$$

où $\boldsymbol{\varepsilon}_l^q \sim \eta(\mathbf{0}_{40}, q\mathbf{I}_{40})$ et \mathcal{M} la résolvante de Eq. (3.64) sur Δt .

Le premier ensemble a priori est engendré par M perturbations gaussiennes normalisées de $\mathbf{x}_{0:C}$ (\mathbf{x}_0 est obtenu après une intégration de $5 \times 10^3 \times \Delta t$ d'un point aléatoire). Les vecteurs d'observations sont engendrés à chaque pas de temps Δt avec $\mathcal{H}_l = \mathbf{I}_n$ (système complètement observé) et $\mathbf{R}_l = \mathbf{I}_n$. Le pas des différences finies est fixé à $\epsilon = 10^{-4}$; le critère d'arrêt de Gauss-Newton est $\Delta = 10^{-3}$; le nombre d'itérations maximal de Gauss-Newton est $j_{\max} = 10$. La moyenne temporelle de la RMSE est calculée avec :

$$\text{RMSE} = \frac{1}{c\sqrt{n}} \sum_{k=0}^{c-1} \left\| \mathbf{x}_{kS+L-\text{lag}} - \mathbf{x}_{kS+L-\text{lag}}^a \right\|, \quad (3.65)$$

où c est le nombre total de cycles. Avec $\text{lag} = 0$ on mesure la RMSE de filtrage ; alors que pour $\text{lag} = L$ on mesure la RMSE de lissage. L'inflation multiplicative est choisie empiriquement parmi

$$\{1; 1, 02; 1, 05; 1, 1; 1, 15; 1, 2; 1, 25; 1, 3; 1, 4; 1, 5; 1, 75; 2; 2, 5; 3; 4\}, \quad (3.66)$$

de façon à maximiser la RMSE de filtrage. La version bundle et des rotations aléatoires ont été appliquées à tous les algorithmes.

3.5.1 Validation de l'IEnKS-Q

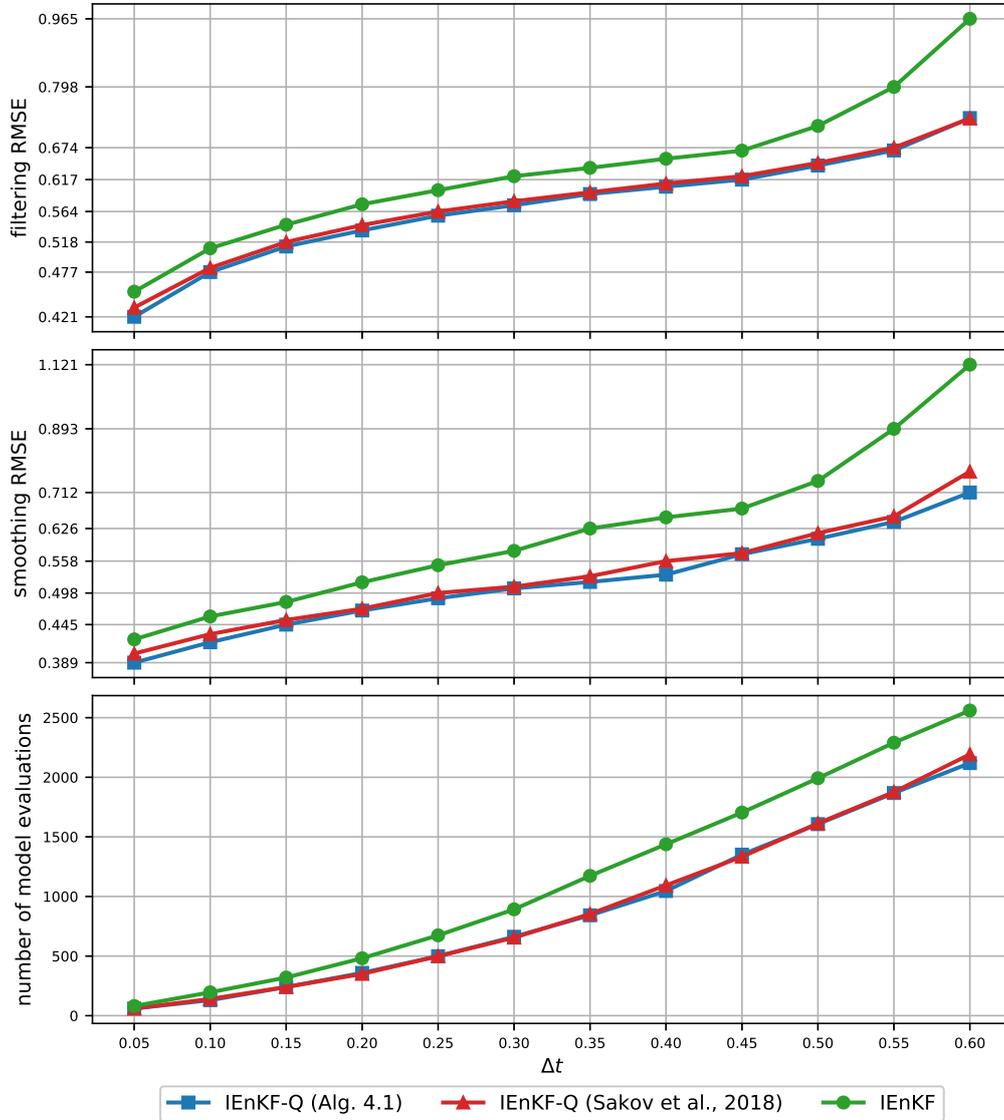


FIGURE 3.3 – RMSE de lissage et de filtrage et nombre moyen d'évaluation du modèle de l'IEnKF-Q défini dans Sakov et al. (2018), de l'IEnKF-Q défini dans l'algorithme 3.1 ($G = 0$, variantes 3.2,3.3,3.4,3.5) et de l'IEnKF fortement contraint, en fonction de Δt .

La figure 3.3 est similaire à la figure 1 de Sakov et al. (2018). La RMSE de filtrage, de lissage et le nombre moyen d'évaluations du modèle sur δt sont tracés en fonction de Δt pour l'IEnKF-Q de Sakov et al. (2018), l'IEnKF-Q de l'algorithme 3.1 ($G = 0$, $L = S = 1$ avec les variantes 3.2, 3.3, 3.4, 3.5) et l'IEnKF fortement contraint (i.e. où l'erreur modèle n'est pas prise en compte autrement que par l'inflation). Le paramètre Δt règle la non

linéarité du système. L'intensité d'erreur modèle est $q = 0.2 \times \Delta t$. Le nombre total de vecteurs d'observations assimilés est 10^4 . Les ensembles de l'IEEnKF ont $M = 20$ membres ; les ensembles de l'IEEnKF-Q Sakov et al. (2018) ont $M = 61$ membres divisés en $M^b = 20$ membres pour l'ébauche (suffisamment pour recouvrir les sous-espaces neutres et instables, Bocquet and Carrassi, 2017) et $M^a = 41$ membres pour l'erreur modèle (suffisamment pour que la matrice d'anomalie soit un facteur exact de la matrice de covariance d'erreur modèle tout en ayant le vecteur $\mathbf{1}_{40}$ dans son noyau) ; L'IEEnKF-Q algorithme 3.1 a une taille effective d'ensembles de $m = 59$ divisé en $m^b = 19$ et $m^a = 40$. Avec ces paramètres, les algorithmes sont équivalents en termes de paramètres de fenêtre et de nombre de membres des ensembles.

Les RMSE augmentent avec Δt puisque l'approximation Gaussienne des densités a priori est de moins en moins valide. Le nombre d'itérations de Gauss-Newton augmente aussi (non affiché) parce que la fonction de coût est de moins en moins quadratique et donc plus difficile à minimiser. Ainsi, un cycle d'assimilation demande plus de temps et requiert plus d'évaluations du modèle. L'IEEnKS fortement contraint a les pires RMSE puisqu'il ne tient pas compte de l'erreur modèle ; bien que ses performances soient relativement bonnes grâce à l'inflation qui les compense. Les RMSE des deux versions de l'IEEnKF-Q sont proches.

3.5.2 Test du paramètre G

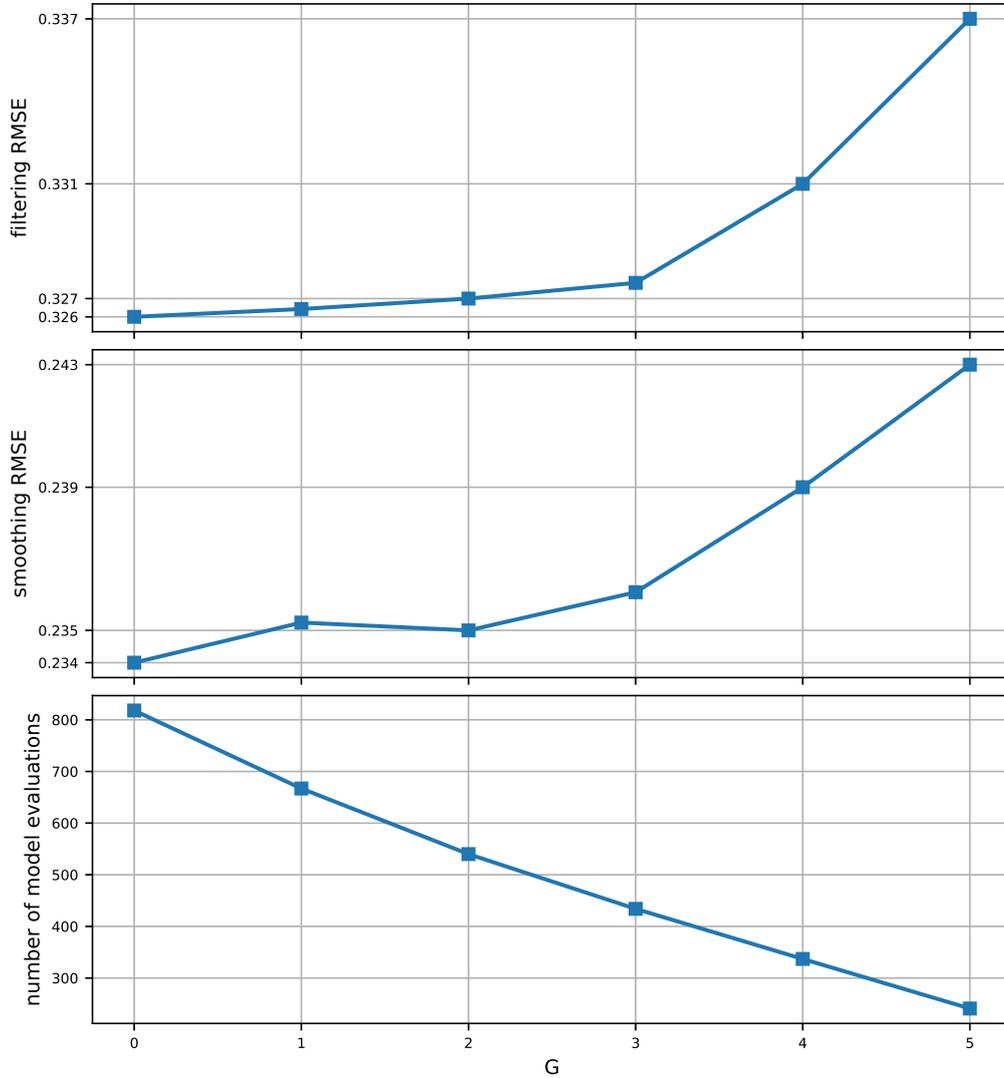


FIGURE 3.4 – RMSE de lissage et de filtrage, nombre moyen d'évaluations du modèle de l'IEnKS-Q ($L = S = 5$, variantes 3.2, 3.3, 3.4, 3.5) en fonction de G .

La figure 3.4 affiche les RMSE de filtrage et de lissage, le nombre moyen d'évaluations du modèle sur δt en fonction de G de l'IEnKS-Q ($L = S = 5$, variants 3.2, 3.3, 3.4, 3.5). L'intensité d'erreur modèle est $q = 10^{-2}$; l'intervalle temporel entre observations est $\Delta t = 0.05$; le nombre total de vecteurs d'observations assimilés est 5×10^4 . La dimension effective de l'ensemble est $m = m^b + S \times m^a$ avec $m^b = (G + 1) \times 40$ et $m^a = 40$. Ainsi, les matrices de déviation sont carrées pour ne pas avoir d'erreurs liées à la réduction

de dimension. Les RMSE augmentent avec G puisque les approximations gaussienne des densités a priori s'étendent sur plus de pas de temps. En d'autres termes, la perte d'information est plus conséquente lors de la propagation. Le nombre d'évaluations du modèle diminue avec G parce qu'elles sont repoussées de l'analyse à la propagation où elles ne sont pas multipliées par le nombre d'itérations de Gauss-Newton. En effet, chaque itération de Gauss-Newton demande une propagation de t_G à t_L ; alors que l'étape de propagation demande une propagation de t_L à t_{G+L} .

3.5.3 Test avec des modèles de substitution

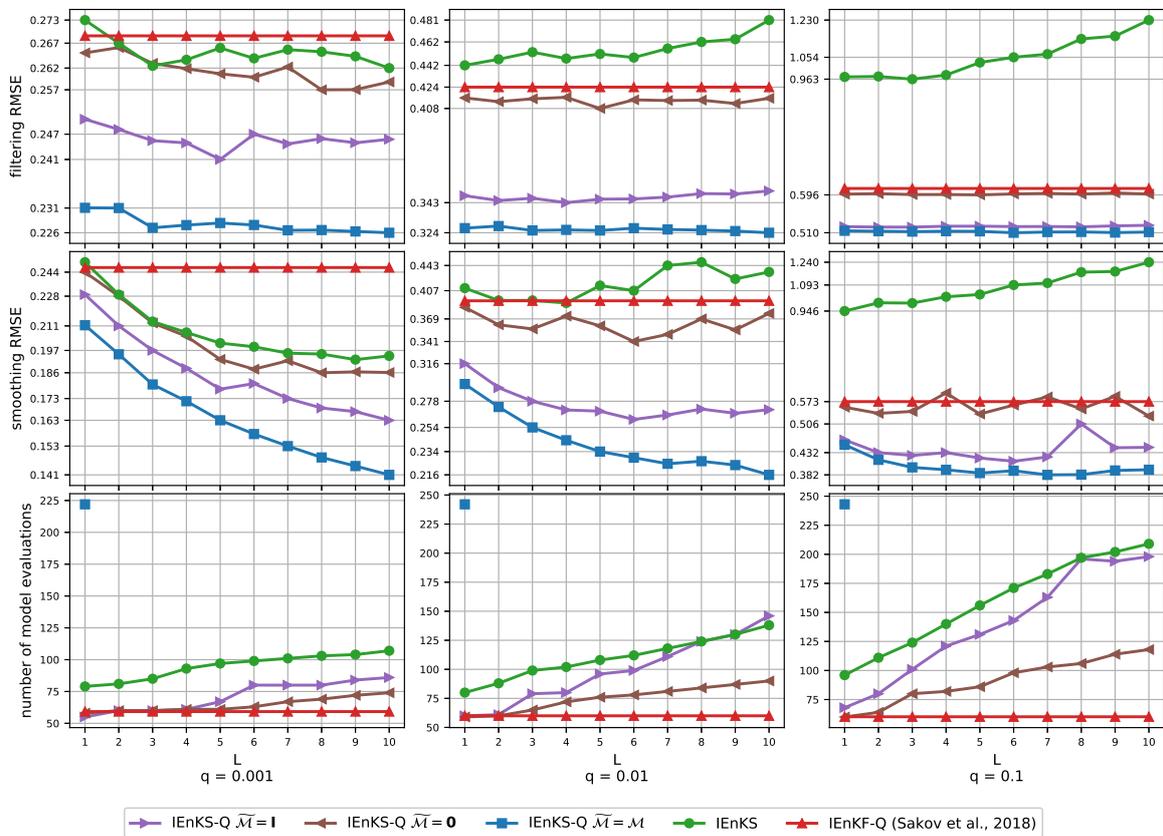


FIGURE 3.5 – RMSE de filtrage et de lissage, nombre moyen d'évaluations du modèle de l'IEEnKS-Q ($G = 0$, variantes 3.2, 3.3, 3.4, 3.6) avec différents modèles de substitution, de l'IEEnKF-Q Sakov et al. (2018) et de l'IEEnKS fortement contraint en fonction de L , $S = L$ et q .

La figure 3.5 affiche les RMSE de lissage et de filtrage, le nombre moyen d'évaluations du modèle (limité à 250) en fonction de L , $S = L$ et de q de l'IEEnKF-Q Sakov et al. (2018)

en tant que référence, de l'IEEnKS fortement contraint et de l'IEEnKS-Q ($G = 0$, variantes 3.2, 3.3, 3.4, 3.6) avec deux modèles de substitution simples $\tilde{\mathcal{M}} = \mathbf{0}_{40 \times 40}$, \mathbf{I}_{40} et l'évaluation complète du modèle $\tilde{\mathcal{M}} = \mathcal{M}$. La condition $S = L$ signifie que les fenêtres d'assimilation ne se superposent pas. Les intensités d'erreur modèle sont $q = 10^{-3}, 10^{-2}, 10^{-1}$. L'intervalle temporel entre les observations est $\Delta t = 0.05$. Le nombre total de vecteurs d'observation assimilés est 10^4 . L'IEEnKF-Q et l'IEEnKS ont les mêmes dimension d'ensemble que dans la figure 3.3; L'IEEnKS-Q a les mêmes dimension d'ensemble que dans la figure 3.4. Les matrices de déviation sont décomposées avec une SVD limitée, les premières $m^0 = 19$ directions associées aux plus grandes valeurs singulières engendrent un ensemble de $M^0 = 20$ membres évalué par le modèle. Les autres directions engendrent un ensemble évalué par le modèle de substitution. L'IEEnKF-Q ne dépend pas de L par définition. Les différences en nombre d'évaluations du modèle des algorithmes sont dues aux différences en nombre d'itérations de Gauss-Newton. Avec une petite intensité d'erreur modèle ($q = 10^{-3}$), l'IEEnKS est quelquefois plus performant que l'IEEnKF-Q du fait de l'augmentation de L améliore les performances. L'IEEnKS-Q sans modèle de substitution ($\tilde{\mathcal{M}} = \mathcal{M}$) affiche les meilleures RMSE mais son nombre d'évaluations du modèle peut être prohibitif. En outre, l'IEEnKS-Q avec modèle de substitution $\tilde{\mathcal{M}} = \mathbf{I}_{40}$ affiche les meilleures performances. Il nécessite aussi moins d'inflation (non affiché) parce que la variance dans les directions associées à de faibles valeurs singulières ne sont pas totalement supprimées comme dans le cas $\tilde{\mathcal{M}} = \mathbf{0}_{40 \times 40}$. L'IEEnKS-Q gère donc efficacement l'erreur modèle et ne demande pas plus d'évaluations du modèle que l'IEEnKS. L'amélioration des performances avec la longueur de fenêtre L est plus forte lorsque l'erreur modèle est de faible intensité. Ceci est dû au fait que l'information contenue par des observations éloignées dans le temps est bruitée par l'erreur modèle. Quand l'erreur modèle est de forte intensité, les RMSE de l'IEEnKS-Q sont pratiquement constantes avec L . Ceci suggère que la non linéarité du modèle est négligeable devant l'erreur modèle. Dans ce cas, il n'est pas justifié d'utiliser de longues fenêtres d'assimilation.

3.6 Conclusion

Une méthode 4D ensemble variationnelle faiblement contrainte a été proposée. Elle généralise l'IEEnKS fortement contraint Bocquet and Sakov (2014) en tenant compte de l'erreur modèle et généralise l'IEEnKF-Q Sakov et al. (2018) à différentes fenêtres d'assimilation ($L > 1, K > 1$). Sa première itération de Gauss-Newton, peut être vue comme équivalente à l'AEnKF-Q développé dans Sakov and Bocquet (2018). L'IEEnKS-Q vise à tenir compte de la non linéarité dans de longues fenêtres d'assimilation et de

l'erreur modèle additive tout en gardant de bonnes performances de lissage et filtrage.

Un autre schéma de méthode 4D ensemble variationnelle faiblement contrainte a été proposé dans [Amezcu et al. \(2017\)](#) où la variable de contrôle est le premier état de la fenêtre avec le biais d'erreur modèle. Nous avons montré que cette approche conduisait à plus d'approximations gaussiennes des densités a priori et négligeais les covariances croisées. Ainsi, l'intrication entre l'erreur modèle et la dynamique non linéaire est négligée. Ainsi, lorsque l'opérateur d'observation est linéaire, la minimisation de la fonction de coût se réduit à un IEnKS fortement contraint avec une covariance d'erreur d'observation corrigée.

L'IEnKS-Q démarre avec une approximation de la moyenne et de la covariance de $p(\mathbf{v}|\mathbf{y}_{:L-S})$ disponible par récurrence. À l'aide de ces statistiques, la variable d'ébauche \mathbf{v} est paramétrisée dans l'espace de l'ensemble par \mathbf{w} . Le changement de variable correspondant est fait pour normaliser ces statistiques et préconditionner la fonction de coût a priori. La règle de Bayes et une approximation gaussienne de $p(\mathbf{w}|\mathbf{y}_{:L-S})$ nous donnent ensuite une expression analytique de la fonction de coût a posteriori $J(\mathbf{w}|\mathbf{y}_{:L})$. L'analyse variationnelle consiste alors à linéariser itérativement les opérateurs $(\mathcal{F}_l)_{l=K,\dots,L}$ à la manière d'une méthode de Gauss-Newton. Ces linéarisations sont obtenues récursivement avec une interpolation sur des ensembles. Ainsi, au terme de cette analyse variationnelle, on dispose d'une approximation de la moyenne et de la covariance à posteriori dont on peut se servir pour engendrer un ensemble a posteriori déterministe échantillonnant $p(\mathbf{w}|\mathbf{y}_{:L})$. Lors de la propagation, les statistiques de la densité a priori du prochain cycle $p(\mathbf{v}'|\mathbf{y}_{:L})$ sont calculées, éventuellement avec des linéarisations du modèle provenant de l'analyse. Finalement, la marginalisation réduit à leur dimension initiale de façon optimale les statistiques a priori du prochain cycle. Une éventuelle structure bloc diagonale de la matrice de covariance peut être exploitée. Cela assure l'hypothèse de récurrence du prochain cycle et rend l'algorithme séquentiel.

Cependant, la dimension de la variable de contrôle doit augmenter avec la longueur de fenêtre pour décrire précisément les statistiques de la trajectoire des états. Malheureusement, le nombre d'évaluations du modèle est proportionnel au carré de la longueur de fenêtre, ce qui peut devenir coûteux avec des modèles complexes de grande dimension. Pour résoudre ce problème, les matrices de déviation sont décomposées pour séparer le nombre de membres de l'ensemble d'avec le nombre d'évaluations du modèle.

Nos expériences numériques montrent que pour $L = 1$, l'algorithme 3.1 et l'IEnKF-Q de [Sakov and Bocquet \(2018\)](#) ont des résultats proches, ce qui conforte le fait que l'IEnKS-Q est sa généralisation. Quand $L > 1$, les RMSE augmentent avec G suggérant que l'approximation gaussienne de la densité a priori est moins valable pour la variable

d'état. Les RMSE diminuent aussi avec L mais seulement lorsque l'erreur modèle n'est pas trop forte.

La décomposition des matrices de déviation montre que l'on peut efficacement séparer le nombre d'évaluations du modèle d'avec le nombre de membres de l'ensemble et que même des modèles de substitution triviaux améliorent sensiblement les RMSE.

Cette étude fournit un cadre général pour la construction de méthodes d'assimilation de données 4D ensemble variationnelles faiblement contraintes. La suite de ce travail pourra consister à lancer plus d'expériences numériques avec des matrices de covariance non diagonales ou avec différentes répartitions d'observations. Aussi, la localisation de la méthode devra être développée pour tester l'efficacité de notre implémentation avec des modèles de plus haute dimension. Cela continuera certainement le travail de [Sakov et al. \(2018\)](#) pour l'IEEnKF-Q. Un potentiel lien entre les décompositions de matrices de déviation et la localisation pourrait être examiné. Dans [Bishop et al. \(2017\)](#); [Bocquet \(2016\)](#); [Farchi and Bocquet \(2019\)](#), un ensemble de M états est mis à jour depuis un ensemble augmenté de taille bien plus grande dans un contexte de localisation de matrices de covariance avec un filtre de Kalman d'ensemble « square root ».

Chapitre 4

Conclusion

Nous avons vu que l'enjeu de l'assimilation de données est l'approximation de la densité a posteriori $p(\mathbf{x}|\mathbf{y})$ car l'expression analytique fournie par la règle de Bayes est généralement difficile à exploiter. Dans un contexte géophysique, la complexité des modèles et la taille conséquente des jeux de données imposent à cette approximation de ne retenir qu'une quantité limitée de données. On peut alors choisir d'approcher la densité a posteriori par une densité gaussienne car cette dernière est caractérisée par sa moyenne et matrice de covariance. La meilleure candidate gaussienne au sens de l'entropie relative est la densité gaussienne égalisant moyenne et matrice de covariance :

$$p(\mathbf{x}|\mathbf{y}) \simeq \mathcal{N}\left(\mathbf{x}|\mathbb{E}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}], \mathbb{C}_{p(\mathbf{x}|\mathbf{y})}[\mathbf{x}]\right). \quad (4.1)$$

Ce choix est conforté par le fait que la moyenne a posteriori constitue le meilleur estimateur de \mathbf{x} fonction de \mathbf{y} en erreur quadratique moyenne. Malheureusement, moyenne et matrice de covariance a posteriori ne peuvent pas être calculées algébriquement. Une alternative consiste à faire une approximation plus grossière en approchant la densité jointe par la densité gaussienne de même moyenne et matrice de covariance :

$$p(\mathbf{x}, \mathbf{y}) \simeq \mathcal{N}\left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \middle| \mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \mathbb{C}_{p(\mathbf{x}, \mathbf{y})} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}\right). \quad (4.2)$$

Dans le cadre de cette approximation, moyenne et matrice de covariance a posteriori peuvent être calculées algébriquement à partir de la moyenne et de la matrice de covariance de la densité jointe. L'approximation de la moyenne a posteriori qui en résulte est alors le meilleur estimateur de \mathbf{x} fonction affine de \mathbf{y} en erreur quadratique moyenne. Cependant, ce caractère affine peut être considéré comme restrictif. L'estimateur de maximum a posteriori, quant à lui, se calcule à l'aide des techniques itératives d'optimi-

sation numérique et constitue un estimateur non linéaire. Cette estimation par maximum a posteriori est néanmoins incomplète car elle ne donne pas, sans davantage de développements, d'estimateur de second ordre et donc pas d'approximation gaussienne de la densité a posteriori (ce qui empêchera plus tard le cyclage de la méthode). On peut cependant garder l'idée de ces méthodes itératives : un modèle de la fonction de coût est calculé, le problème est résolu pour ce modèle, sa solution permet ensuite d'améliorer le modèle, etc. Dans un cadre probabiliste, un modèle gaussien de la densité a posteriori découle d'une approximation affine de l'opérateur d'observation.

Avec la dimension temporelle, l'assimilation peut devenir séquentielle. Chaque cycle se décompose alors en une étape d'analyse qui calcule la densité a posteriori à partir de la densité a priori et une étape de propagation qui calcule la densité a priori du prochain cycle en décalant la fenêtre d'assimilation dans le temps. Le filtre particulaire utilise une représentation ensembliste des densités. L'analyse se fait par échantillonnage d'importance et la propagation est une application membre à membre du modèle. Le filtre de Kalman utilise une représentation gaussienne des densités avec des hypothèses de linéarité des opérateurs. L'analyse consiste à calculer algébriquement la moyenne et la matrice de covariance a posteriori. Puisque le modèle est une fonction affine, la propagation consiste à calculer la moyenne et matrice de covariance d'une transformation affine de la densité a posteriori. Le filtre de Kalman d'ensemble utilise des représentations ensemblistes et gaussiennes des densités. L'analyse utilise une approximation gaussienne de la densité jointe pour estimer les moments a posteriori. Les membres de l'ensemble a posteriori sont ensuite engendrés tels que leurs moyenne et matrice de covariance égalisent ces moments a posteriori. Pour le filtre de Kalman d'ensemble déterministe, l'ensemble a posteriori est engendré tel que sa moyenne et matrice de covariance empiriques soient égales à ces moments a posteriori. Le 4DVar consiste à calculer le maximum a posteriori en minimisant itérativement la fonction de coût a posteriori. Bien que cette fonction de coût s'étende sur plusieurs pas de temps, la méthode n'est pas pleinement séquentielle car aucun estimateur de second ordre n'est fourni.

L'assimilation de données variationnelle d'ensemble propose de résoudre ce problème afin de combiner la précision de l'analyse variationnelle avec une estimation des statistiques allant jusqu'au second ordre. La difficulté de produire de tels algorithmes provient des natures locales et globales de ces méthodes. L'IEnKS combine ces approches en donnant une portée statistique aux approximations affines des opérateurs. La performance de l'assimilation ensemble variationnelle est donc liée à la convergence des méthodes d'optimisation itératives. L'IEnKS utilise cette analyse variationnelle pour minimiser une fonction de coût 4D (i.e. s'étalant sur plusieurs pas de temps) dans un cadre séquen-

tiel. Ceci permet différentes stratégies d'estimation de la même densité. Ces stratégies ne sont pas équivalentes lorsque les opérateurs sont non linéaires.

En effet, l'IEEnKS utilise à chaque cycle une approximation gaussienne de la densité a priori. Dans le cadre d'opérateurs non linéaires, de l'information est donc perdue à chaque cycle de la méthode. Or, pour assimiler le même nombre d'observations, les méthodes utilisant de grandes fenêtres d'assimilation nécessitent moins de cycles. Elles reposent donc moins fréquemment sur l'approximation gaussienne de la densité a priori. Leurs performances doivent donc être meilleures. Ceci est numériquement vérifié jusqu'à une certaine longueur maximale de fenêtre. Au delà, le bassin d'attraction du minimum global de la fonction de coût est trop petit pour contenir l'itéré initial. L'analyse variationnelle échoue et conduit à une divergence rapide de l'algorithme. L'analyse quasi-statistique repousse cette limite en ajoutant graduellement les observations à la fonction de coût lors de minimisations répétées. Le succès de la méthode tient au fait que l'assimilation d'une observation réduit la variance d'erreur sur l'état. Le minimiseur est alors statistiquement plus proche de l'état. La taille du bassin d'attraction du minimum global peut alors être réduite en ajoutant de nouvelles observations. L'IEEnKS quasi-statistique (IEEnKS_{QS}) proposé dans cette thèse est donc en mesure d'utiliser de grandes fenêtres d'assimilation pour atteindre de meilleures performances. Cependant, le coût de calcul de la méthode est élevé. Il peut être réduit en limitant le nombre d'itérations des minimisations intermédiaires.

Les résultats précédents reposent sur le caractère fortement contraint du système d'assimilation. En effet, le quasi-statisme gère la non linéarité du modèle dans ses plus petits détails. Si ces détails sont négligeables devant l'erreur modèle, il ne sert à rien d'en tenir compte. Pour améliorer les performances de l'IEEnKS il faut alors en développer une version tenant mathématiquement compte de l'erreur modèle. L'IEEnKS-Q proposé dans cette thèse est une méthode 4D ensemble variationnelle faiblement contrainte qui généralise l'IEEnKS fortement contraint (Bocquet and Sakov, 2014) en tenant compte de l'erreur modèle et généralise l'IEEnKF-Q (Sakov et al., 2018) à différentes fenêtres d'assimilation ($L > 1, K > 1$). Sa première itération de Gauss-Newton, peut être vue comme équivalente à l'AEnKF-Q développé dans Sakov and Bocquet (2018). L'IEEnKS-Q tient ainsi compte de la non linéarité dans de longues fenêtres d'assimilation et de l'erreur modèle additive. Pour cela, l'IEEnKS-Q commence par changer la variable de la trajectoire de l'état en y incorporant des variables d'erreur modèle pour que la nouvelle variable de contrôle soit plus gaussienne. La moyenne et la covariance a priori de cette variable d'ébauche sont ensuite utilisées pour la paramétriser par la variable d'ensemble. Ce changement de variable normalise ses statistiques et préconditionne la fonction de

coût a priori. L'approximation gaussienne de la densité a priori de la variable d'ensemble permet, avec la règle de Bayes, d'obtenir une expression simplifiée de la fonction de coût a posteriori de l'IEnKS-Q. L'analyse variationnelle consiste alors à linéariser itérativement les opérateurs de cette fonction de coût à la manière d'une méthode de Gauss-Newton. Ces linéarisations sont obtenues récursivement avec une interpolation sur des ensembles. Au terme de cette analyse variationnelle, on dispose d'une approximation de la moyenne et de la covariance a posteriori. Ces moments servent ensuite à engendrer un ensemble a posteriori déterministe échantillonnant la variable d'ensemble a posteriori. Cet ensemble est ensuite mis à jour pour produire un échantillon de la trajectoire a posteriori ainsi que de la variable d'ébauche du prochain cycle. Finalement, la marginalisation réduit à leur dimension initiale et de façon optimale les statistiques a priori du prochain cycle. Cela assure l'hypothèse de récurrence du prochain cycle et rend l'algorithme séquentiel. Cependant, la dimension de la variable de contrôle doit augmenter avec la longueur de fenêtre pour décrire précisément les statistiques de la trajectoire des états. Ce qui induit un nombre d'évaluations du modèle proportionnel au carré de la longueur de fenêtre. Pour résoudre ce problème, les matrices de déviation sont décomposées pour séparer le nombre de membres de l'ensemble d'avec le nombre d'évaluations du modèle.

Les algorithmes de l'IEnKS quasi-statique et de l'IEnKS-Q n'ont été testés que sur des modèles de Lorenz de faible dimension. Une perspective opérationnelle consisterait à tester ces méthodes avec différents modèles plus réalistes ou de plus grande dimension. D'autres expériences numériques avec différentes répartitions ou densités d'observations pourront aussi être menées. La localisation de l'IEnKS-Q sera sûrement nécessaire.

Nous avons aussi vu dans cette thèse que l'objectif de l'assimilation de données est l'approximation de la densité a posteriori $p(\mathbf{x}|\mathbf{y})$ et pas seulement l'estimation de \mathbf{x} par une fonction de \mathbf{y} . Dans ce cadre, l'analyse variationnelle ne doit donc pas se contenter de calculer itérativement une estimation de \mathbf{x} , elle doit calculer itérativement une approximation de $p(\mathbf{x}|\mathbf{y})$. C'est ce que fait l'IEnKS en fournissant, à chaque itération de son analyse, une approximation de la moyenne et de la matrice de covariance a posteriori. La fonction de coût réellement minimisée par l'IEnKS doit donc être plus générale et porter sur ces deux moments. Une perspective intéressante serait alors de préciser le lien entre l'analyse variationnelle de l'IEnKS et la minimisation itérative de l'entropie relative entre la densité a posteriori et une densité gaussienne paramétrée par sa moyenne et sa matrice de covariance :

$$\mathcal{G}(\mathbf{x}, \mathbf{A}) = \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[-\ln \frac{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})}{p(\mathbf{x}|\mathbf{y})} \right], \quad (4.3)$$

dont la solution est $\mathbf{x}^* = \mathbb{E}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}]$, $\mathbf{A}^* = \mathbb{C}_{p(\mathbf{x}|\mathbf{y})} [\mathbf{x}]$ d'après l'annexe A.2. Cependant, l'évaluation d'une telle fonction de coût est problématique puisque l'intégration se fait par rapport à la densité $p(\mathbf{x}|\mathbf{y})$. Une alternative peut être de « renverser » l'entropie relative pour obtenir la fonction de coût :

$$\mathcal{J}(\mathbf{x}, \mathbf{A}) = \mathbb{E}_{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})} \left[-\ln \frac{p(\mathbf{x}|\mathbf{y})}{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})} \right], \quad (4.4)$$

qui se simplifie en :

$$\mathcal{J}(\mathbf{x}, \mathbf{A}) \propto \mathbb{E}_{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})} [G(\mathbf{x}|\mathbf{y})] - \frac{1}{2} \ln |\mathbf{A}|. \quad (4.5)$$

Le terme $\mathbb{E}_{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})} [G(\mathbf{x}|\mathbf{y})]$ peut être vu comme une évaluation de $G(\mathbf{x}|\mathbf{y})$ en $\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})$ au sens des distributions alors que $-\frac{1}{2} \ln |\mathbf{A}|$ peut être vu comme un terme de pénalisation/régularisation. Avec une hypothèse de domination, les dérivées de $\mathcal{J}(\mathbf{x}, \mathbf{A})$ selon (\mathbf{x}, \mathbf{A}) ne font intervenir que les dérivées de $\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})$ selon (\mathbf{x}, \mathbf{A}) . Les dérivées de $\mathcal{J}(\mathbf{x}, \mathbf{A})$ peuvent donc être estimées à l'aide d'évaluations en $G(\mathbf{x}|\mathbf{y})$ d'un échantillon (stochastique ou déterministe) de $\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})$. La fonction de coût $\mathcal{J}(\mathbf{x}, \mathbf{A})$ peut ainsi être minimisée avec, par exemple, une méthode de Newton sans avoir à évaluer les dérivées de $G(\mathbf{x}|\mathbf{y})$. Si les erreurs d'échantillonnage sont négligées, la méthode est alors localement convergente, on ne sait cependant pas vers quoi. Un lien peut quand même être établi entre dérivées de $\mathcal{J}(\mathbf{x}, \mathbf{A})$ et dérivées de $G(\mathbf{x}|\mathbf{y})$ avec :

$$\mathcal{J}(\mathbf{x}, \mathbf{A}) = \frac{1}{2} \ln |\mathbf{A}| + \mathbb{E}_{\mathcal{N}(\mathbf{w}|\mathbf{0}, \mathbf{I})} \left[G\left(\mathbf{x} = \mathbf{x} + \mathbf{A}^{\frac{1}{2}} \mathbf{w} | \mathbf{y}\right) \right]. \quad (4.6)$$

On aura par exemple :

$$\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}, \mathbf{A}) = \mathbb{E}_{\mathcal{N}(\mathbf{w}|\mathbf{0}, \mathbf{I})} \left[\nabla_{\mathbf{x}} G\left(\mathbf{x} = \mathbf{x} + \mathbf{A}^{\frac{1}{2}} \mathbf{w} | \mathbf{y}\right) \right], \quad (4.7a)$$

$$= \mathbb{E}_{\mathcal{N}(\mathbf{x}|\mathbf{x}, \mathbf{A})} \left[\nabla_{\mathbf{x}} G(\mathbf{x}|\mathbf{y}) \right], \quad (4.7b)$$

ce qui nous conforte dans l'idée que $\mathcal{J}(\mathbf{x}, \mathbf{A})$ est une fonction de coût plus générale découlant d'une autre forme d'évaluation de $G(\mathbf{x}|\mathbf{y})$.

References

- Amezcuca, J., Goodliff, M., and van Leeuwen, P.-J. (2017). A weak-constraint 4densem-blevar part i : formulation and simple model experiments. *Tellus A*, 69 :1271564.
- Anderson, B. D. O. and Moore, J. B. (1979). *Optimal Filtering*. Dover Publications.
- Asch, M., Bocquet, M., and Nodet, M. (2017). *Data Assimilation : Methods, Algorithms and Applications*.
- Auligné, T., Ménétrier, B., Lorenc, A. C., and Buehner, M. (2016). Ensemble variational integrated localized data assimilation. *Monthly Weather Review*.
- Bain, A. and Crisan, D. (2009). *Fundamentals of Stochastic Filtering*. Springer-Verlag New York.
- Bengtsson, L., Ghil, M., and Källén, E. (1981). *Dynamic Meteorology : Data Assimilation Methods*. Springer-Verlag New York.
- Bernardo, J. M. and Smith, A. F. (2008). Bayesian theory. *Bayesian Theory*, pages 1–595.
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer-Verlag New York.
- Bishop, C., Whitaker, J., and Lei, L. (2017). Gain form of the ensemble transform kalman filter and its relevance to satellite data assimilation with model space ensemble covariance localization. *Monthly Weather Review*, 145 :4575–4592.
- Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform kalman filter. part i : Theoretical aspects. *Monthly Weather Review*.
- Björck, A. (1996). *Numerical methods for least squares problems*. Society for Industrial and Applied Mathematics.

- Bocquet, M. (2011). Ensemble kalman filtering without the intrinsic need for inflation. *Nonlinear Processes in Geophysics*, 18(5) :735–750.
- Bocquet, M. (2016). Localization and the iterative ensemble kalman smoother. *Q. J. R. Meteorol. Soc.*, 142 :1075–1089.
- Bocquet, M. and Carrassi, A. (2017). Four-dimensional ensemble variational data assimilation and the unstable subspace. *Tellus A*, 69 :1304504.
- Bocquet, M., Raanes, P. N., and Hannart, A. (2015). Expanding the validity of the ensemble kalman filter without the intrinsic need for inflation. *Nonlin. Processes Geophys.*, 22 :645–662.
- Bocquet, M. and Sakov, P. (2012). Combining inflation-free and iterative ensemble kalman filters for strongly nonlinear systems. *Nonlinear Processes in Geophysics*.
- Bocquet, M. and Sakov, P. (2013). Joint state and parameter estimation with an iterative ensemble kalman smoother. *Nonlin. Processes Geophys.*, 20 :803–818.
- Bocquet, M. and Sakov, P. (2014). An iterative ensemble kalman smoother. *Quarterly Journal of the Royal Meteorological Society*, 140(682) :1521–1535.
- Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B. (2010). Inter-comparison of variational data assimilation and the ensemble kalman filter for global deterministic nwp. part i : Description and single-observation experiments. *Monthly Weather Review*.
- Burgers, G., Jan van Leeuwen, P., and Evensen, G. (1998). Analysis scheme in the ensemble kalman filter. *Monthly Weather Review*.
- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2017a). Data assimilation in the geosciences - an overview on methods, issues and perspectives. *Wiley Interdisciplinary Reviews : Climate Change*.
- Carrassi, A., Bocquet, M., Hannart, A., and Ghil, M. (2017b). Estimating model evidence using data assimilation. *Q. J. R. Meteorol. Soc.*, 143 :866–880.
- Cohn, S. E. (1997). An introduction to estimation theory. *Data Assimilation Office Note Series on Global Modeling and Data Assimilation*.
- Cohn, S. E., Sivakumaran, N. S., and Todling, R. (1994). A fixed-lag kalman smoother for retrospective data assimilation. *Monthly Weather Review*.

- Cosme, E., Verron, J., Brasseur, P., Blum, J., and Auroux, D. (2012). Smoothing problems in a bayesian framework and their linear gaussian solutions. *Mon. Wea. Rev.*, 140 :683–695.
- Cover, T. M. and Thomas, J. A. (2005). *Elements of Information Theory*. Elements of Information Theory.
- Dennis, J. and Schnabel, R. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. SIAM J. Optim.
- Di Persio, L. (2018). Asymptotic expansions of integrals : Statistical mechanics and quantum theory. Technical report.
- Dimet, F. X. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. *Tellus A*.
- Doucet, A. and Johansen, A. M. (2011). A tutorial on particle filtering and smoothing : Fiteen years later. *The Oxford handbook of nonlinear filtering*.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*.
- Evensen, G. (2003). The ensemble kalman filter : Theoretical formulation and practical implementation. *Ocean Dynamics*.
- Evensen, G. (2009). *Data Assimilation : The Ensemble Kalman Filter*. Springer-Verlag Berlin Heidelberg.
- Farchi, A. and Bocquet, M. (2019). On the efficiency of covariance localisation of the EnKF using augmented ensembles. *Front. Appl. Math. Stat.* Accepted for publication.
- Feller, W. (2008). An introduction to probability theory and its applications. *Second edition*.
- Fertig, E. J., Harlim, J., and Hunt, B. R. (2007). A comparative study of 4d-var and a 4d ensemble kalman filter : Perfect model simulations with lorenz-96. *Tellus, Series A : Dynamic Meteorology and Oceanography*.
- Fillion, A., Bocquet, M., and Gratton, S. (2018). Quasi-static ensemble variational data assimilation : a theoretical and numerical study with the iterative ensemble kalman smoother. *Nonlinear Processes in Geophysics*, 25 :315–334.

- Gallouët, T. and Herbin, R. (2013). *Mesure, intégration, probabilités*. Ellipses.
- Goodliff, M., Amezcua, J., and Van Leeuwen, P. J. (2015). Comparing hybrid data assimilation methods on the lorenz 1963 model with increasing non-linearity. *Tellus A*, 67 :26928.
- Gordon, N., Salmond, D., and Smith, A. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F Radar and Signal Processing*.
- Hamill, T. M. and Snyder, C. (2000). A hybrid ensemble kalman filter 3d variational analysis scheme. *Monthly Weather Review*.
- Hamill, T. M., Whitaker, J. S., and Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Review*.
- Haussaire, J.-M. and Bocquet, M. (2015). A low-order coupled chemistry meteorology model for testing online and offline data assimilation schemes. *Geoscientific Model Development Discussions*, 8 :7347–7394.
- Houtekamer, Herschel L.P., L. and Mitchell (1998). Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review*, 126(3) :796–811.
- Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*.
- Hunt, B. R., Kalnay, E., Kostelich, E. J., Ott, E., Patil, D. J., Sauer, T., Szunyogh, I., Yorke, J. a., and Zimin, a. V. (2004). Four-dimensional ensemble kalman filtering. *Tellus A*.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos : A local ensemble transform kalman filter. *Physica D : Nonlinear Phenomena*.
- Jardak, M. and Talagrand, O. (2018). Ensemble variational assimilation as a probabilistic estimator. part i : The linear and weak non-linear case. *Nonlinear Processes in Geophysics Discussions*, pages 1–39.
- Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Elsevier Science.

- Johnson, O. (2004). *Information Theory and the Central Limit Theorem*. Imperial College Press.
- Judd, K., Smith, L., and Weisheimer, A. (2004). Gradient free descent : shadowing, and state estimation using limited derivative information. *Physica D*, 190 :153–166.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1) :35.
- Kong, A. and Liu, J. S. (1994). Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*.
- Liu, C., Xiao, Q., and Wang, B. (2008). An ensemble-based four-dimensional variational data assimilation scheme. part i : Technical formulation and preliminary test. *Monthly Weather Review*.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20 :130–141.
- Lorenz, E. N. and Emanuel, K. A. (1998). Optimal sites for supplementary weather observations : Simulation with a small model. *J. Atmos. Sci.*, 55 :399–414.
- Mandel, J., Bergou, E., Gürol, S., Gratton, S., and Kusanický, I. (2016). Hybrid levenberg-marquardt and weak-constraint ensemble kalman smoother method. *Non-linear Processes in Geophysics*.
- Miller, R. N., Ghil, M., and Gauthiez, F. (1994). Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, 51 :1037–1056.
- Nocedal, J. and Wright, S. J. (2006). *Numerical optimization*. Springer series in operations research.
- Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D., and Yorke, J. A. (2004). A local ensemble kalman filter for atmospheric data assimilation. *Tellus A : Dynamic Meteorology and Oceanography*.
- Pires, C., Vautard, R., and Talagrand, O. (1996). On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus A*, 48 :96–121.
- Puntanen, S. and Styan, G. P. H. (2011). Best linear unbiased estimation in linear models. In Lovric, M., editor, *International Encyclopedia of Statistical Science*, pages 141–144. Springer Berlin Heidelberg, Berlin, Heidelberg.

- Raynaud, L., Berre, L., and Desroziers, G. (2009). Objective filtering of ensemble-based background-error variances. *Quarterly Journal of the Royal Meteorological Society*, 135(642) :1177–1199.
- Robert, C., Blayo, E., and Verron, J. (2006). Reduced-order 4d-var : A preconditioner for the incremental 4d-var data assimilation method. *Geophysical Research Letters*.
- Robert, C., Durbiano, S., Blayo, E., Verron, J., Blum, J., and Le Dimet, F. X. (2005). A reduced-order strategy for 4d-var data assimilation. *Journal of Marine Systems*.
- Robert, C. P. and Casella, G. (1999). Monte carlo statistical methods. *Springer Texts in Statistics*.
- Sakov, P. and Bertino, L. (2011). Relation between two common localisation methods for the enkf. *Computational Geosciences*.
- Sakov, P. and Bocquet, M. (2018). Asynchronous data assimilation with the enkf in presence of additive model error. *Tellus A*, 70 :1414545.
- Sakov, P., Haussaire, J.-M., and Bocquet, M. (2018). An iterative ensemble kalman filter in presence of additive model error. *Q. J. R. Meteorol. Soc.*, 0 :0.
- Sakov, P. and Oke, P. R. (2008). Implications of the form of the ensemble transformation in the ensemble square root filters. *Monthly Weather Review*.
- Sakov, P., Oliver, D. S., and Bertino, L. (2012). An iterative enkf for strongly nonlinear systems. *Monthly Weather Review*, 140 :1988–2004.
- Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*.
- Stoer, J. and Bulirsch, R. (1982). *Introduction to Numerical Analysis*. Springer.
- Swanson, K., Vautard, R., and Pires, C. (1998). Four-dimensional variational assimilation and predictability in a quasi-geostrophic model. *Tellus A*, 50 :369–390.
- Talagrand, O. (1997). Assimilation of observations, an introduction. *Journal of the Meteorological Society of Japan*.
- Tarantola, A. (2004). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.

-
- Trémolet, Y. (2006). Accounting for an imperfect model in 4d-var. *Q. J. R. Meteorol. Soc.*, 132 :2483–2504.
- Trevisan, A., D’Isidoro, M., and Talagrand, O. (2010). Four-dimensional variational assimilation in the unstable subspace and the optimal subspace dimension. *Q. J. R. Meteorol. Soc.*, 136 :487–496.
- Tukey, J. W. and Wiener, N. (1952). The extrapolation, interpolation and smoothing of stationary time series with engineering applications. *Journal of the American Statistical Association*.
- Walters, P. (1982). *An Introduction to Ergodic Theory*. Graduate texts in mathematics. Springer-Verlag New York.
- Wikle, C. K. and Berliner, L. M. (2007). A bayesian tutorial for data assimilation. *Physica D : Nonlinear Phenomena*.
- Ye, J., Rey, D., Kadakia, N., Eldridge, M., Morone, U. I., Rozdeba, P., Abarbanel, H. D. I., and Quinn, J. C. (2015). Systematic variational method for statistical nonlinear state and parameter estimation. *Phys. Rev. E*, 92 :52901.

Annexe A

Annexe de l'introduction

A.1 Maximum d'entropie

Montrons que la densité gaussienne $q(\mathbf{x}) = \mathcal{N}(\mathbf{x}^b, \mathbf{B})$ est l'unique maximiseur de l'entropie :

$$\mathbb{H}_{r(\mathbf{x})}[r(\mathbf{x})] = \mathbb{E}_{r(\mathbf{x})}[-\ln r(\mathbf{x})], \quad (\text{A.1})$$

parmi toutes les densités $r(\mathbf{x})$ de moyenne \mathbf{x}^b et de covariance \mathbf{B} . Soit $r(\mathbf{x}) \neq q(\mathbf{x})$ une densité de moyenne \mathbf{x}^b et de covariance \mathbf{B} , on a :

$$\mathbb{H}_{r(\mathbf{x})}[r(\mathbf{x})] = \mathbb{H}_{r(\mathbf{x})}\left[\frac{r(\mathbf{x})}{q(\mathbf{x})}\right] + \mathbb{H}_{r(\mathbf{x})}[q(\mathbf{x})], \quad (\text{A.2})$$

Or $\mathbb{H}_{r(\mathbf{x})}\left[\frac{q(\mathbf{x})}{r(\mathbf{x})}\right]$ est l'entropie relative entre les densités $r(\mathbf{x})$ et $q(\mathbf{x})$ (e.g., [Bishop, 2006](#); [Cover and Thomas, 2005](#)). Elle est strictement positive en vertu de l'inégalité de Jensen et de la stricte convexité de $x \mapsto \ln x$:

$$\mathbb{H}_{r(\mathbf{x})}\left[\frac{q(\mathbf{x})}{r(\mathbf{x})}\right] = \mathbb{E}_{r(\mathbf{x})}\left[-\ln \frac{q(\mathbf{x})}{r(\mathbf{x})}\right], \quad (\text{A.3a})$$

$$> -\ln \mathbb{E}_{r(\mathbf{x})}\left[\frac{q(\mathbf{x})}{r(\mathbf{x})}\right], \quad (\text{A.3b})$$

$$> 0. \quad (\text{A.3c})$$

De plus, puisque $q(\mathbf{x})$ est gaussienne, $\mathbf{x} \mapsto \ln[q(\mathbf{x} = \mathbf{x})]$ est une fonction quadratique, le calcul de $\mathbb{H}_{r(\mathbf{x})}[q(\mathbf{x})] = \mathbb{E}_{r(\mathbf{x})}[-\ln q(\mathbf{x})]$ ne fait donc intervenir que la moyenne et la covariance de $r(\mathbf{x})$ qui sont égales à celles de $q(\mathbf{x})$ par hypothèse. Donc $\mathbb{H}_{r(\mathbf{x})}[q(\mathbf{x})] =$

$\mathbb{H}_{q(\mathbf{x})} [q(\mathbf{x})]$ et :

$$\mathbb{H}_{r(\mathbf{x})} [r(\mathbf{x})] < \mathbb{H}_{q(\mathbf{x})} [q(\mathbf{x})]. \quad (\text{A.4})$$

A.2 Minimum d'entropie relative

Soit $p(\mathbf{x})$ une densité de moyenne \mathbf{x}^a et de covariance \mathbf{A} . Montrons que la gaussienne $q(\mathbf{x}) = \mathcal{N}(\mathbf{x}^a, \mathbf{A})$ minimise l'entropie relative (cf A.1) :

$$\mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{p(\mathbf{x})} \right] \geq 0, \quad (\text{A.5})$$

parmi toutes les densités $r(\mathbf{x})$ gaussiennes. On a :

$$\mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{p(\mathbf{x})} \right] = \mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{q(\mathbf{x})} \right] + \mathbb{H}_{p(\mathbf{x})} \left[\frac{q(\mathbf{x})}{p(\mathbf{x})} \right]. \quad (\text{A.6})$$

Puisque $\mathbf{x} \mapsto \ln \left[\frac{r(\mathbf{x})}{q(\mathbf{x})} \right]$ est une fonction quadratique, le calcul de $\mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{q(\mathbf{x})} \right]$ ne fait intervenir que la moyenne et la covariance de $p(\mathbf{x})$ qui sont égales à celles de $q(\mathbf{x})$ par hypothèse. Donc

$$\mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{q(\mathbf{x})} \right] = \mathbb{H}_{q(\mathbf{x})} \left[\frac{r(\mathbf{x})}{q(\mathbf{x})} \right], \quad (\text{A.7a})$$

$$\geq 0. \quad (\text{A.7b})$$

Finalement,

$$\mathbb{H}_{p(\mathbf{x})} \left[\frac{r(\mathbf{x})}{p(\mathbf{x})} \right] \geq \mathbb{H}_{p(\mathbf{x})} \left[\frac{q(\mathbf{x})}{p(\mathbf{x})} \right]. \quad (\text{A.8})$$

A.3 Meilleur estimateur en erreur quadratique moyenne

Soit \mathbf{x}, \mathbf{y} un couple de variables aléatoires à valeurs dans $\mathbb{R}^n \times \mathbb{R}^d$, F l'espace vectoriel des fonctions de \mathbb{R}^n dans \mathbb{R}^d et E un sous-espace vectoriel de F . Soit $\mathcal{H} \in E$ un estimateur et $\boldsymbol{\varepsilon} = \mathbf{y} - \mathcal{H}(\mathbf{x})$ l'erreur d'estimation. On a équivalence entre les propositions :

1. \mathcal{H} est un meilleur estimateur en erreur quadratique de E
2. Pour tout $\mathcal{F} \in E$, $\mathbb{E}_{p(\mathbf{x}, \boldsymbol{\varepsilon})} \langle \boldsymbol{\varepsilon} | \mathcal{F}(\mathbf{x}) \rangle = 0$

Géométriquement, l'optimum correspond ici à une orthogonalité. Ceci est illustré Fig. (A.1).

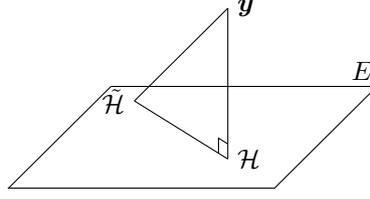


FIGURE A.1 – Optimum et orthogonalité

\Rightarrow : Soit $\mathcal{F} \in E$ et $\alpha \neq 0$, posons $\tilde{\mathcal{H}} = \mathcal{H} + \alpha\mathcal{F}$. Alors $\tilde{\mathcal{H}} \in E$ et d'après 1 :

$$\mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \left\| \mathbf{y} - \tilde{\mathcal{H}}(\mathbf{x}) \right\|^2 = \mathbb{E}_{p(\mathbf{x}, \varepsilon)} \left\| \varepsilon - \alpha\mathcal{F}(\mathbf{x}) \right\|^2, \quad (\text{A.9a})$$

$$= \mathbb{E}_{p(\varepsilon)} \left\| \varepsilon \right\|^2 + \alpha^2 \mathbb{E}_{p(\mathbf{x})} \left\| \mathcal{F}(\mathbf{x}) \right\|^2, \quad (\text{A.9b})$$

$$- 2\alpha \mathbb{E}_{p(\mathbf{x}, \varepsilon)} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle, \quad (\text{A.9c})$$

$$\geq \mathbb{E}_{p(\varepsilon)} \left\| \varepsilon \right\|^2. \quad (\text{A.9d})$$

Donc

$$\alpha \mathbb{E}_{p(\mathbf{x}, \varepsilon)} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle \leq \frac{\alpha^2}{2} \mathbb{E}_{p(\mathbf{x})} \left\| \mathcal{F}(\mathbf{x}) \right\|^2. \quad (\text{A.10})$$

L'énoncé précédent est aussi vrai avec la substitution $\alpha \leftarrow -\alpha$, ce qui donne :

$$\left| \mathbb{E}_{p(\mathbf{x}, \varepsilon)} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle \right| \leq \frac{|\alpha|}{2} \mathbb{E}_{p(\mathbf{x})} \left\| \mathcal{F}(\mathbf{x}) \right\|^2. \quad (\text{A.11})$$

Finalement $\mathbb{E}_{p(\mathbf{x}, \varepsilon)} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle = 0$.

\Leftarrow : Soit $\tilde{\mathcal{H}} \in E$, posons $\mathcal{F} = \tilde{\mathcal{H}} - \mathcal{H}$. Alors $\mathcal{F} \in E$ et d'après 2 :

$$\mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \left\| \mathbf{y} - \tilde{\mathcal{H}}(\mathbf{x}) \right\|^2 = \mathbb{E}_{p(\mathbf{x}, \varepsilon)} \left\| \varepsilon - \mathcal{F}(\mathbf{x}) \right\|^2, \quad (\text{A.12a})$$

$$= \mathbb{E}_{p(\varepsilon)} \left\| \varepsilon \right\|^2 + \mathbb{E}_{p(\mathbf{x})} \left\| \mathcal{F}(\mathbf{x}) \right\|^2, \quad (\text{A.12b})$$

$$- 2\mathbb{E}_{p(\mathbf{x}, \varepsilon)} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle, \quad (\text{A.12c})$$

$$\geq \mathbb{E}_{p(\mathbf{x}, \mathbf{y})} \left\| \mathbf{y} - \mathcal{H}(\mathbf{x}) \right\|^2. \quad (\text{A.12d})$$

A.3.1 Avec E le sous-espace des fonctions affines

Si E est l'ensemble des fonctions affines, posons :

$$\mathcal{H}(\mathbf{x}) = \mathbb{E}_{p(\mathbf{y})}[\mathbf{y}] + \mathbb{C}_{p(\mathbf{y}, \mathbf{x})}[\mathbf{y}, \mathbf{x}] \mathbb{C}_{p(\mathbf{x})}[\mathbf{x}]^{-1} \left(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})}[\mathbf{x}] \right), \quad (\text{A.13a})$$

$$\varepsilon = \mathbf{y} - \mathcal{H}(\mathbf{x}). \quad (\text{A.13b})$$

Alors $\mathcal{H} \in E$ et $\mathbb{E}_{p(\varepsilon)}[\varepsilon] = \mathbf{0}$, $\mathbb{C}_{p(\varepsilon, \mathbf{x})}[\varepsilon, \mathbf{x}] = \mathbf{0}$. Soit $\mathcal{F} : \mathbf{x} \mapsto \mathbf{f} + \mathbf{F}(\mathbf{x} - \mathbb{E}_{p(\mathbf{x})}[\mathbf{x}]) \in E$ on a :

$$\begin{aligned} \mathbb{E}_{p(\varepsilon, \mathbf{x})} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle &= \langle \mathbb{E}_{p(\varepsilon)}[\varepsilon] | \mathbf{f} \rangle + \text{tr}(\mathbf{F}\mathbf{C}[\mathbf{x}, \varepsilon]), \\ &= 0. \end{aligned}$$

Donc \mathcal{H} est le meilleur estimateur affine en erreur quadratique moyenne de E .

A.3.2 Avec E l'espace des fonctions

Si $E = F$, posons :

$$\mathcal{H}(\mathbf{x}) = \mathbb{E}_{p(\mathbf{y}|\mathbf{x})}[\mathbf{y}], \quad (\text{A.15a})$$

$$\varepsilon = \mathbf{y} - \mathcal{H}(\mathbf{x}). \quad (\text{A.15b})$$

Alors $\mathcal{H} \in E$ et $\mathbb{E}_{p(\varepsilon|\mathbf{x})}[\varepsilon] = \mathbf{0}$. Soit $\mathcal{F} \in E$, on a d'après le théorème de l'espérance totale :

$$\mathbb{E}_{p(\varepsilon, \mathbf{x})} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle = \mathbb{E}_{p(\mathbf{x})} \mathbb{E}_{p(\varepsilon|\mathbf{x})} \langle \varepsilon | \mathcal{F}(\mathbf{x}) \rangle, \quad (\text{A.16a})$$

$$= \mathbb{E}_{p(\mathbf{x})} \langle \mathbb{E}_{p(\varepsilon|\mathbf{x})}[\varepsilon] | \mathcal{F}(\mathbf{x}) \rangle, \quad (\text{A.16b})$$

$$= 0. \quad (\text{A.16c})$$

Donc \mathcal{H} est le meilleur estimateur en erreur quadratique moyenne de F .

A.4 Variable d'ensemble et erreur d'échantillonnage

Soit \mathbf{x} une variable aléatoire telle que :

$$\mathbb{E}_{p(\mathbf{x})}[\mathbf{x}] = \mathbf{0}, \quad (\text{A.17a})$$

$$\mathbb{C}_{p(\mathbf{x})}[\mathbf{x}] = (\mathbf{X}^b) (\mathbf{X}^b)^T. \quad (\text{A.17b})$$

Par décomposition orthogonale de \mathbf{x} sur $\text{Im}(\mathbf{X}^b) \oplus \text{Im}(\mathbf{X}^b)^\perp$, il existe un unique couple $(\mathbf{v}, \mathbf{u}) \in \text{Im}(\mathbf{X}^b) \times \text{Im}(\mathbf{X}^b)^\perp$ tel que :

$$\mathbf{x} = \mathbf{v} + \mathbf{u}. \quad (\text{A.18})$$

Montrons que $\mathbf{u} = \mathbf{0}$ au sens de :

$$\mathbb{E}_{p(x)} [\mathbf{u}] = \mathbf{0}, \quad (\text{A.19a})$$

$$\mathbb{V}_{p(x)} [\mathbf{u}] = \mathbf{0}. \quad (\text{A.19b})$$

En prenant l'espérance de la décomposition Eq. (A.18) on obtient :

$$\mathbf{0} = \mathbb{E}_{p(x)} [\mathbf{u}] + \mathbb{E}_{p(x)} [\mathbf{v}], \quad (\text{A.20})$$

donc $\mathbb{E}_{p(x)} [\mathbf{u}] = \mathbb{E}_{p(x)} [\mathbf{v}] = \mathbf{0}$ par unicité de la décomposition. Soit $\mathbf{a} \in \text{Im}(\mathbf{X}^b)^\perp$ alors :

$$\mathbb{V}_{p(x)} [\mathbf{x}^T \mathbf{a}] = \text{tr}(\mathbf{a}^T \mathbb{C}_{p(x)} [\mathbf{x}] \mathbf{a}), \quad (\text{A.21a})$$

$$= \text{tr}(\mathbf{a}^T (\mathbf{X}^b) (\mathbf{X}^b)^T \mathbf{a}), \quad (\text{A.21b})$$

$$= 0. \quad (\text{A.21c})$$

Par ailleurs,

$$\mathbb{V}_{p(x)} [\mathbf{x}^T \mathbf{a}] = 0 + \mathbb{V}_{p(x)} [\mathbf{v}^T \mathbf{a}], \quad (\text{A.22})$$

donc $\mathbf{v}^T \mathbf{a} = 0$ (presque partout). Puisque cela est vrai pour tout $\mathbf{a} \in \text{Im}(\mathbf{X}^b)^\perp$ et $\mathbf{v} \in \text{Im}(\mathbf{X}^b)^\perp$ alors $\mathbf{v} = \mathbf{0}$ (presque partout).

Annexe B

Annexe de l'assimilation de données ensemble variationnelle quasi-statique

B.1 Performance du 4D-Var dans le cas linéaire et uni-varié

L'objectif de cette annexe est d'établir une relation de récurrence entre la eMSE de chaque cycle du 4D-Var. Avec ces relations, nous obtiendrons l'expression de la eMSE asymptotique du 4D-Var.

Supposons $n = 1$ et $\mathcal{M}^l(x) = \alpha^l x$. Au kème cycle, le gradient et la hessienne de la fonction de coût du 4D-Var Eq. (2.7) sont :

$$\begin{aligned} \nabla J(x_{kS}) = & -\frac{1}{b} (x_{kS}^b - x_{kS}) \\ & - \frac{h}{r} \sum_{l=K}^L \alpha^l (y_{kS+l} - h\alpha^l x_{kS}), \end{aligned} \quad (\text{B.1a})$$

$$\nabla^2 J(x_{kS}) = \frac{1}{b} + \frac{h^2}{r} \sum_{l=K}^L \alpha^{2l}, \quad (\text{B.1b})$$

où la dépendance de la fonction de coût aux observations $y_{kS+L:kS+K}$ n'a pas été notée. On remarque que la hessienne ne dépend ni de $y_{kS+L:kS+K}$, ni de x_{kS} , elle n'est donc pas aléatoire et sera simplement notée $\nabla^2 J$. Puisque les opérateurs sont linéaires, J est une fonction quadratique, définie, positive de x_{kS} . Elle admet donc un unique argminimum

x_{kS}^a qui est caractérisé par l'annulation du gradient :

$$\nabla J(x_{kS}^a) = 0. \quad (\text{B.2})$$

À l'aide d'un développement de Taylor exact autour de l'état x_{kS} , on obtient :

$$0 = \nabla J(x_{kS}) + \nabla^2 J(x_{kS}) \times (x_{kS}^a - x_{kS}). \quad (\text{B.3})$$

En y injectant les Eq. (B.1a,2.1a) on obtient :

$$\nabla^2 J \times (x_{kS}^a - x_{kS}) = \frac{1}{b} (x_{kS}^b - x_{kS}) + \frac{h}{r} \sum_{l=K}^L \alpha^l \varepsilon_{kS+l}. \quad (\text{B.4})$$

La variable aléatoire $x_{kS}^b - x_{kS} = \alpha^S (x_{(k-1)S}^a - x_{(k-1)S})$ est indépendante des erreurs $\varepsilon_{kS+K}, \dots, \varepsilon_{kS+L}$. Ainsi prendre l'espérance du carré de Eq. (B.4) donne l'expression suivante pour la eMSE $P_{kS}^{4D\text{-Var}}$ du 4D-Var au kème cycle :

$$P_{kS}^{4D\text{-Var}} = \mathbb{E} [(x_{kS} - x_{kS}^a)^2], \quad (\text{B.5a})$$

$$= (\nabla J)^{-2} \left(\frac{\alpha^{2S}}{b^2} P_{(k-1)S}^{4D\text{-Var}} + \frac{h^2}{r} \sum_{l=K}^L \alpha^{2l} \right), \quad (\text{B.5b})$$

$$P_{-S}^{4D\text{-Var}} = b. \quad (\text{B.5c})$$

Avec les notations :

$$\Sigma_K^L = \frac{h^2}{r} \sum_{l=K}^L \alpha^{2l} = \frac{h^2}{r} \alpha^{2K} \frac{\alpha^{2S} - 1}{\alpha^2 - 1}, \quad (\text{B.6a})$$

$$\Delta = \frac{\alpha^{2S}}{(1 + b\Sigma_K^L)^2}, \quad (\text{B.6b})$$

on obtient :

$$P_{kS}^{4D\text{-Var}} = \Delta P_{(k-1)S}^{4D\text{-Var}} + \frac{\Sigma_K^L}{\left(\frac{1}{b} + \Sigma_K^L\right)^2}. \quad (\text{B.7})$$

Ainsi, $\{P_{kS}^{4D\text{-Var}}\}_k$ est une suite arithmetico-géométrique. Sa limite $P_{\infty S}^{4D\text{-Var}}$ vérifie :

$$P_{\infty S}^{4D\text{-Var}} = \begin{cases} \infty & \text{si } \Delta \geq 1, \\ \frac{b^2 \Sigma_K^L}{\alpha^{2S}} \frac{\Delta}{1-\Delta} & \text{sinon.} \end{cases} \quad (\text{B.8})$$

La généralisation vers une eMSE asymptotique avec lag $L - l$ est directe :

$$P_{\infty S+l}^{4D\text{-Var}} = \begin{cases} \infty & \text{si } \Delta \geq 1, \\ \frac{b^2 \Sigma_K^L}{\alpha^{2(S-l)}} \frac{\Delta}{1-\Delta} & \text{sinon.} \end{cases} \quad (\text{B.9})$$

Dans le cas multi-varié et diagonal, ces calculs peuvent être conduits indépendamment sur chaque direction. Dans ce cas, la eMSE est simplement la somme des eMSEs uni-variées de chaque direction.

B.2 Performance de l'IEnKS dans le cas linéaire et uni-varié

L'objectif de cette annexe est d'établir une relation de récurrence entre la eMSE de chaque cycle de l'IEnKS. Avec ces relations, nous obtiendrons l'expression de la eMSE asymptotique de l'IEnKS.

Prouvons d'abord par récurrence que pour tout $k \geq 0$, la fonction de coût $G(x_{kS}|y_{K:kS+L})$ est gaussienne de moments :

$$\mathbb{E}_{p(x_{kS}|y_{K:kS+L})}[x_{kS}] = x_{kS}^a, \quad (\text{B.10a})$$

$$\mathbb{V}_{p(x_{kS}|y_{K:kS+L})}[x_{kS}] = (X_{kS}^a)^2, \quad (\text{B.10b})$$

où \mathbb{V} est l'opérateur de variance et x_{kS}^a, X_{kS}^a sont définies par les équations Eq. (2.14a,2.14b). Grâce aux hypothèses Eq. (2.2,2.24c) avec Eq.(2.12), on a :

$$G(x_0|y_{K:L}) \propto J(w_0|y_{K:L}). \quad (\text{B.11})$$

Ainsi, $p(x_0|y_{K:L})$ est gaussienne et ses moments sont donnés par Eq. (B.10a,B.10b) avec $k = 0$. Supposons par récurrence que $p(x_{kS}|y_{K:kS+L})$ est gaussienne et ses moments sont donnés par Eq. (B.10a,B.10b) pour un certain $k \geq 0$. Puisque $x_{(k+1)S} = \mathcal{M}^S(x_{kS})$ et \mathcal{M} est affine, on a :

$$\mathbb{E}_{p(x_{(k+1)S}|y_{K:kS+L})}[x_{(k+1)S}] = \mathcal{M}^S(x_{kS}^a) = x_{(k+1)S}^b, \quad (\text{B.12a})$$

$$\mathbb{V}_{p(x_{(k+1)S}|y_{K:kS+L})}[x_{(k+1)S}] = \alpha^{2S} (X_{kS}^a)^2 = X_{(k+1)S}^b, \quad (\text{B.12b})$$

à l'aide les propriétés de l'espérance et des Eq. (2.24a,2.24b). Ce résultat avec l'Eq. (2.6)

et l'hypothèse Eq. (2.12) permet d'obtenir :

$$G\left(x_{(k+1)S}|y_{K:(k+1)S+L}\right) \propto J\left(w_{(k+1)S+K:(k+1)S+L}\right). \quad (\text{B.13})$$

Ainsi, $p\left(x_{(k+1)S}|y_{K:(k+1)S+L}\right)$ est gaussienne et ses moments sont donnés par Eq. (B.10a,B.10b).

La variance Eq. (B.10b) est reliée à la performance de l'IEnKS par le théorème des espérances totales :

$$P_{kS}^{\text{IEnKS}} = \mathbb{E}_{p(x_{kS}, y_{K:kS+L})} \left[(x_{kS} - x_{kS}^a)^2 \right], \quad (\text{B.14a})$$

$$= \mathbb{E}_{p(y_{K:kS+L})} \mathbb{E}_{p(x_{kS}|y_{K:kS+L})} \left[(x_{kS} - x_{kS}^a)^2 \right], \quad (\text{B.14b})$$

$$= \mathbb{E}_{p(y_{K:kS+L})} \mathbb{V}_{p(x_{kS}|y_{K:kS+L})} [x_{kS}], \quad (\text{B.14c})$$

$$= \mathbb{E}_{p(y_{K:kS+L})} \left[(X_{kS}^a)^2 \right]. \quad (\text{B.14d})$$

Puis avec l'Eq. (2.14b) on obtient les relations de récurrence :

$$\left(X_{(k+1)S}^a\right)^{-2} = \alpha^{-2S} \left(X_{kS}^a\right)^{-2} + \frac{h^2}{r} \sum_{l=K}^L \alpha^{2l}, \quad (\text{B.15a})$$

$$\left(X_0^a\right)^{-2} = b^{-1} + \frac{h^2}{r} \sum_{l=K}^L \alpha^{2l}. \quad (\text{B.15b})$$

Donc X_{kS}^a n'est pas aléatoire et $P_{kS}^{\text{IEnKS}} = (X_{kS}^a)^2$. Les équations (B.15) montrent que la suite des inverses de eMSEs de l'IEnKS est arithmetico-géométrique :

$$\left(P_{kS}^{\text{IEnKS}}\right)^{-1} = \alpha^{-2S} \left(P_{(k-1)S}^{\text{IEnKS}}\right)^{-1} + \Sigma_K^L, \quad (\text{B.16})$$

$$P_{-S}^{\text{IEnKS}} = b^{-1}, \quad (\text{B.17})$$

où la notation Eq. (B.6a) a été utilisée. On obtient ensuite la eMSE asymptotique de l'IEnKS à l'aide des propriétés des suites arithmético-géométriques :

$$P_{\infty S}^{\text{IEnKS}} = \begin{cases} 0 & \text{si } |\alpha| \leq 1, \\ \frac{r}{h^2 \alpha^{2L}} \frac{\alpha^2 - 1}{\alpha^2} & \text{sinon.} \end{cases} \quad (\text{B.18})$$

La généralisation aux eMSEs asymptotiques de lag $L - l$ est directe :

$$P_{\infty S}^{\text{IEnKS}} = \begin{cases} 0 & \text{si } |\alpha| \leq 1, \\ \frac{r}{h^2 \alpha^{2(L-l)}} \frac{\alpha^2 - 1}{\alpha^2} & \text{sinon.} \end{cases} \quad (\text{B.19})$$

Montrons à présent que la eMSE de l'IEnKS est optimale. Soit $\tilde{x}_{kS}^a(y_{K:kS+L})$ l'analyse du 4D-Var ou n'importe quelle autre fonction de $y_{K:kS+L}$. Une décomposition biais-variance (e.g. Bishop, 2006) (e.g. Bishop (2006)) donne :

$$\begin{aligned} \mathbb{E}_{p(x_{kS}, y_{K:kS+L})} \left[(x_{kS} - \tilde{x}_{kS}^a)^2 \right] &= \mathbb{E}_{p(y_{K:kS+L})} \mathbb{V}_{p(x_{kS}|y_{K:kS+L})} [x_{kS}] \\ &\quad + \mathbb{E}_{p(y_{K:kS+L})} \left[\left(\mathbb{E}_{p(x_{kS}|y_{K:kS+L})} [x_{kS}] - \tilde{x}_{kS}^a \right)^2 \right]. \end{aligned} \quad (\text{B.20})$$

En remplaçant les moments avec Eq. (B.10a,B.10b) on obtient :

$$P_{kS}^{\text{4D-Var}} = P_{kS}^{\text{IEnKS}} + \mathbb{E}_{p(y_{K:kS+L})} \left[(x_{kS}^a - \tilde{x}_{kS}^a)^2 \right], \quad (\text{B.21a})$$

$$\geq P_{kS}^{\text{IEnKS}}. \quad (\text{B.21b})$$

Dans le cas multi-varié et diagonal, les calculs peuvent être conduits indépendamment dans chaque directions. Ainsi, la eMSE de ce cas est la somme des eMSEs uni-variées de chaque direction.

B.3 Expression de la fonction de coût moyennée

La fonction de coût moyennée de l'IEnKS $J_{\infty S}$ est la limite quand N tend vers l'infini de :

$$\frac{1}{N} \sum_{k=0}^{N-1} J(\mathbf{w} | \mathbf{y}_{K+kS:kS+L}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2N} \sum_{l=K}^L \sum_{k=0}^{N-1} \left\| \mathbf{y}_{kS+l} - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w}) \right\|_{\mathbf{R}^{-1}}^2. \quad (\text{B.22})$$

En développant le carré de la norme autour de $\mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS})$ avec Eq. (2.1) on obtient :

$$\frac{1}{2} \left\| \mathbf{y}_{kS+l} - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w}) \right\|_{\mathbf{R}^{-1}}^2 = \frac{1}{2} \|\boldsymbol{\varepsilon}_{kS+l}\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \left\| \delta \mathbf{y}_{kS+l} \right\|_{\mathbf{R}^{-1}}^2 + \boldsymbol{\varepsilon}_{kS+l}^T \mathbf{R}^{-1} \delta \mathbf{y}_{kS+l}, \quad (\text{B.23})$$

où $\delta \mathbf{y}_{kS+l} = \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}) - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{kS}^b + \mathbf{X}_{kS}^b \mathbf{w})$. On suppose qu'il existe des variables aléatoires $(\boldsymbol{\varepsilon}_{\infty S}, \mathbf{x}_{\infty S}, \mathbf{x}_{\infty S}^b, \mathbf{X}_{\infty S}^b)$ dont la densité est invariante et ergodique par rapport au décalage temporel :

$$\begin{aligned} T : (\boldsymbol{\varepsilon}_{kS}, \mathbf{x}_{kS}, \mathbf{x}_{kS}^b, \mathbf{X}_{kS}^b) &\mapsto \\ &(\boldsymbol{\varepsilon}_{(k+1)S}, \mathbf{x}_{(k+1)S}, \mathbf{x}_{(k+1)S}^b, \mathbf{X}_{(k+1)S}^b). \end{aligned} \quad (\text{B.24})$$

Puisque les $\{\boldsymbol{\varepsilon}_{kS}\}_k$ sont mutuellement indépendants, indépendants des $\{\mathbf{x}_{kS}, \mathbf{x}_{kS}^b, \mathbf{X}_{kS}^b\}_k$ et identiquement distribués alors $p(\boldsymbol{\varepsilon}_{\infty S}, \mathbf{x}_{\infty S}, \mathbf{x}_{\infty S}^b, \mathbf{X}_{\infty S}^b) = p(\boldsymbol{\varepsilon}_0) p(\mathbf{x}_{\infty S}, \mathbf{x}_{\infty S}^b, \mathbf{X}_{\infty S}^b)$. D'après le théorème ergodique de Birkhoff (Walters (1982)) on obtient :

$$\frac{1}{N} \sum_{k=0}^{N-1} \|\boldsymbol{\varepsilon}_{kS+l}\|_{\mathbf{R}^{-1}}^2 \xrightarrow{N \rightarrow \infty} \mathbb{E}_{p(\boldsymbol{\varepsilon}_0)} \left[\|\boldsymbol{\varepsilon}_0\|_{\mathbf{R}^{-1}}^2 \right] = d, \quad (\text{B.25a})$$

$$\frac{1}{N} \sum_{k=0}^{N-1} \boldsymbol{\varepsilon}_{kS+l}^T \mathbf{R}^{-1} \delta \mathbf{y}_{kS+l} \xrightarrow{N \rightarrow \infty} \mathbb{E}_{p(\boldsymbol{\varepsilon}_0)} [\boldsymbol{\varepsilon}_0]^T \mathbb{E}_{p(\delta \mathbf{y}_{\infty S+l})} [\delta \mathbf{y}_{\infty S+l}] = 0, \quad (\text{B.25b})$$

$$\frac{1}{N} \sum_{k=0}^{N-1} \|\delta \mathbf{y}_{kS+l}\|_{\mathbf{R}^{-1}}^2 \xrightarrow{N \rightarrow \infty} \mathbb{E}_{p(\delta \mathbf{y}_{\infty S+l})} \left[\|\delta \mathbf{y}_{\infty S+l}\|_{\mathbf{R}^{-1}}^2 \right], \quad (\text{B.25c})$$

où $\delta \mathbf{y}_{\infty S+l} = \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{\infty S}) - \mathcal{H} \circ \mathcal{M}^l(\mathbf{x}_{\infty S}^b + \mathbf{X}_{\infty S}^b \mathbf{w})$. Finalement,

$$J_{\infty S}(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 + \frac{dS}{2} + \frac{1}{2} \sum_{l=K}^L \mathbb{E}_{p(\delta \mathbf{y}_{\infty S+l})} \left[\|\delta \mathbf{y}_{\infty S+l}\|_{\mathbf{R}^{-1}}^2 \right]. \quad (\text{B.26})$$

Annexe C

Annexe du lisseur de Kalman itératif faiblement contraint

C.1 Incompatibilité entre les approches « biais d'erreur modèle » et « forçage d'erreur modèle »

Avec les hypothèses Eq. (3.17), la variable $\mathbf{x}_1 = \mathcal{M}_1(\mathbf{x}_0) + \boldsymbol{\beta}_1$ est non gaussienne à cause de la non linéarité du modèle. Alors qu'avec les hypothèses Eq. (3.18), \mathbf{x}_1 est gaussienne tant que $G > 0$. Ces deux ensembles d'hypothèses sont donc incompatibles si $G > 0$ et le modèle est non linéaire.

Si $G = 0$ et $L > 1$, $\boldsymbol{\beta}_2$ est gaussien avec les hypothèses Eq. (3.17). Alors qu'avec les hypothèses Eq. (3.18), $\boldsymbol{\beta}_2$ est généralement une fonction non linéaire de variables aléatoires gaussiennes :

$$\boldsymbol{\beta}_2 = \mathcal{M}_2(\mathcal{M}_1(\mathbf{x}_0) + \boldsymbol{\varepsilon}_1^q) + \boldsymbol{\varepsilon}_2^q - \mathcal{M}_{2 \leftarrow 0}(\mathbf{x}_0),$$

elle n'est donc généralement pas gaussienne. Les deux ensembles d'hypothèses sont alors incompatibles si $G = 0, L > 1$ et le modèle est non linéaire.

Si $G = 0$ et $L = 1$ alors $\boldsymbol{\beta}_1 = \boldsymbol{\varepsilon}_1^q$. Les deux ensembles d'hypothèses sont donc compatibles.

C.2 Réduction de la fonction de coût 4DVar faiblement contrainte avec le biais d'erreur modèle comme variable de contrôle

Une supposition centrale faite dans [Amezcuca et al. \(2017\)](#) est que les opérateurs d'observation sont affines :

$$\mathcal{H}_l(\mathbf{x} + \delta\mathbf{x}) = \mathcal{H}_l(\mathbf{x}) + \mathbf{H}_l\delta\mathbf{x}. \quad (\text{C.1})$$

Dans ce cas, la fonction de coût Eq. (3.16) est quadratique selon $\beta_{1:L}$. En effet, si nous définissons :

$$\mathcal{D}(\mathbf{x}_0) = \text{vec}(\mathbf{y}_1 - \mathcal{H}_1 \circ \mathcal{M}_{1 \leftarrow 0}(\mathbf{x}_0), \dots, \mathbf{y}_L - \mathcal{H}_L \circ \mathcal{M}_{L \leftarrow 0}(\mathbf{x}_0)), \quad (\text{C.2a})$$

$$\beta = \text{vec}(\beta_1, \dots, \beta_L), \quad (\text{C.2b})$$

$$\mathbf{y} = \text{vec}(\mathbf{y}_1, \dots, \mathbf{y}_L), \quad (\text{C.2c})$$

$$\mathbf{Q}^\beta = \text{diag}(\mathbf{Q}_1^\beta, \dots, \mathbf{Q}_L^\beta), \quad (\text{C.2d})$$

$$\mathbf{R} = \text{diag}(\mathbf{R}_1, \dots, \mathbf{R}_L), \quad (\text{C.2e})$$

$$\mathbf{H} = \text{diag}(\mathbf{H}_1, \dots, \mathbf{H}_L), \quad (\text{C.2f})$$

alors Eq. (3.16) peut être écrite de manière compacte :

$$J(\mathbf{x}_0, \beta; \mathbf{y}) = \frac{1}{2} \|\mathcal{D}(\mathbf{x}_0) - \mathbf{H}\beta\|_{\mathbf{R}^{-1}}^2 + \frac{1}{2} \|\beta\|_{(\mathbf{Q}^\beta)^{-1}}^2 + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0^{-1}}^2. \quad (\text{C.3})$$

Cette fonction de coût peut être interprétée comme l'opposé du logarithme de la densité a posteriori :

$$J(\mathbf{x}_0, \beta; \mathbf{y}) \propto -\ln p(\mathbf{x}_0, \beta | \mathbf{y}), \quad (\text{C.4})$$

si les hypothèses Eq. (3.17) sont imposées. Cette densité a posteriori et cette fonction de coût peuvent être factorisées et décomposées en :

$$p(\mathbf{x}_0, \beta | \mathbf{y}) = p(\beta | \mathbf{x}_0, \mathbf{y}) p(\mathbf{x}_0 | \mathbf{y}), \quad (\text{C.5a})$$

$$J(\mathbf{x}_0, \beta; \mathbf{y}) = \frac{1}{2} \|\beta - \beta^*\|_{\nabla_\beta^2 J}^2 + J(\mathbf{x}_0, \beta^*, \mathbf{y}), \quad (\text{C.5b})$$

où β^* est le minimiseur de la fonction de coût en β . Ces décompositions suggèrent, suivant [Sakov et al. \(2018\)](#), une procédure de deux étapes pour minimiser Eq. (3.16) ou, de manière équivalente, maximiser Eq. (C.5a).

Dans la première étape, $p(\mathbf{x}_0|\mathbf{y})$ est maximisé selon \mathbf{x}_0 . Sa plus simple expression est obtenue par marginalisation de $p(\mathbf{x}_0, \boldsymbol{\beta}|\mathbf{y})$ selon $\boldsymbol{\beta}$:

$$-\ln p(\mathbf{x}_0|\mathbf{y}) \propto \frac{1}{2} \|\mathcal{D}(\mathbf{x}_0)\|_{(\mathbf{R}+\mathbf{H}\mathbf{Q}\mathbf{H}^T)^{-1}}^2 + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}_0^{-1}}^2. \quad (\text{C.6})$$

C'est la fonction de coût d'un IEnKS [Bocquet and Sakov \(2014\)](#) fortement contraint avec une matrice de covariance d'erreur d'observation \mathbf{R} redéfinie en $\mathbf{R} + \mathbf{H}\mathbf{Q}\mathbf{H}^T$.

Dans la seconde étape, $p(\boldsymbol{\beta}|\mathbf{x}_0, \mathbf{y})$ est maximisé selon $\boldsymbol{\beta}$ sachant \mathbf{x}_0 obtenu lors de la première étape. On a :

$$-\ln p(\boldsymbol{\beta}|\mathbf{x}_0, \mathbf{y}) \propto \frac{1}{2} \left\| \boldsymbol{\beta} - \mathbf{Q}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{Q}\mathbf{H}^T)^{-1} \mathcal{D}(\mathbf{x}_0) \right\|_{\mathbf{Q}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}}^2. \quad (\text{C.7})$$

Cette expression est quadratique selon $\boldsymbol{\beta}$ et son minimiseur a une expression analytique pouvant être directement lue.

Cette procédure de minimisation en deux étapes généralise la preuve donnée en section 3 de [Sakov et al. \(2018\)](#) pour l'IEnKF-Q à un lisseur de fenêtre d'assimilation de longueur L .