

# Transport optimal pour l'assimilation de données images

Nelson Feyeux

► **To cite this version:**

Nelson Feyeux. Transport optimal pour l'assimilation de données images. Analyse numérique [math.NA]. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAM076 . tel-01480695v2

**HAL Id: tel-01480695**

**<https://hal.inria.fr/tel-01480695v2>**

Submitted on 11 Jan 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE la Communauté UNIVERSITÉ**  
**GRENOBLE ALPES**

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 Août 2006

Présentée par

**Nelson FEYEUX**

Thèse dirigée par **Arthur VIDARD**  
et codirigée par **Maëlle NODET**

préparée au sein **du Laboratoire Jean Kuntzmann**  
et de **l'École Doctorale Mathématiques, Sciences et Technologies**  
**de l'Information, Informatique**

# Transport optimal pour l'assimilation de données images.

Thèse soutenue publiquement le **08/12/2016**,  
devant le jury composé de :

**M. Emmanuel MAITRE**

Professeur, Grenoble-INP, Président

**Mme Julie DELON**

Professeure, Université Paris Descartes, Rapporteur

**M. Didier AUROUX**

Professeur, Université Nice-Sophia-Antipolis, Rapporteur

**Mme Mélanie ROCHOUX**

Chargée de recherche, CERFACS, Examinatrice

**M. Yann BRENIER**

Directeur de recherche, École Polytechnique, Examineur

**M. Nicolas PAPADAKIS**

Chargé de recherche, CNRS, Examineur

**M. Arthur VIDARD**

Chargé de recherche, Inria, Directeur de thèse

**Mme Maëlle NODET**

Maître de conférences, Université Grenoble-Alpes, Co-Directrice de thèse



## Résumé

Pour prédire l'évolution d'un système physique, nous avons besoin d'initialiser le modèle mathématique le représentant, donc d'estimer la valeur de l'état du système au temps initial. Cet état n'est généralement pas directement mesurable car souvent trop complexe. En revanche, nous disposons d'informations du système, prises à des temps différents, incomplètes, mais aussi entachées d'erreurs, telles des observations, de précédentes estimations, etc. Combiner ces différentes informations partielles et imparfaites pour estimer la valeur de l'état fait appel à des méthodes d'assimilation de données dont l'idée est de trouver un état initial proche de toutes les informations. Ces méthodes sont très utilisées en météorologie.

Nous nous intéressons dans cette thèse à l'assimilation d'images, images qui sont de plus en plus utilisées en tant qu'observations. La spécificité de ces images est leur cohérence spatiale, l'œil humain peut en effet percevoir des structures dans les images que les méthodes classiques d'assimilation ne considèrent généralement pas. Elles ne tiennent compte que des valeurs de chaque pixel, ce qui résulte dans certains cas à des problèmes d'amplitude dans l'état initial estimé. Pour résoudre ce problème, nous proposons de changer d'espace de représentation des données : nous plaçons les données dans un espace de Wasserstein où la position des différentes structures compte. Cet espace, équipé d'une distance de Wasserstein, est issue de la théorie du transport optimal et trouve beaucoup d'applications en imagerie notamment.

Dans ce travail nous proposons une méthode d'assimilation variationnelle de données basée sur cette distance de Wasserstein. Nous la présentons ici, ainsi que les algorithmes numériques liés et des expériences sur des modèles de Saint-Venant 1D montrant ses spécificités. Nous verrons dans les résultats comment elle permet de corriger ce qu'on appelle erreurs de position.

**Mots-clés :** assimilation de données, transport optimal, distance de Wasserstein, problèmes inverses, traitement d'images.

## Abstract

Forecasting of a physical system is computed by the help of a mathematical model. This model needs to be initialized by the state of the system at initial time. But this state is not directly measurable and data assimilation techniques are generally used to estimate it. They combine all sources of information such as observations (that may be sparse in time and space and potentially include errors), previous forecasts, the model equations and error statistics. The main idea of data assimilation techniques is to find an initial state accounting for the different sources of informations. Such techniques are widely used in meteorology, where data and particularly images are more and more numerous due to the increasing number of satellites and other sources of measurements. This, coupled with developments of meteorological models, have led to an ever-increasing quality of the forecast.

Spatial consistency is one specificity of images. For example, human eyes are able to notice structures in an image. However, classical methods of data assimilation do not handle such structures because they take only into account the values of each pixel separately. In some cases it leads to a bad initial condition. To tackle this problem, we proposed to change the representation of an image : images are considered here as elements of the Wasserstein space endowed with the Wasserstein distance coming from the optimal transport theory. In this space, what matters is the positions of the different structures.

This thesis presents a data assimilation technique based on this Wasserstein distance. This technique and its numerical procedure are first described, then experiments are carried out with a 1D Shallow-Water model, and results are shown. In particular, it appears that this technique was able to give an analysis of corrected position.

**Keywords :** data assimilation, optimal transport, Wasserstein distance, inverse problems, image processing.

## Remerciements

J'ai ressenti ces trois années de thèse comme étant une expérience unique.

Je voudrais spécialement remercier mes directeurs de thèse, Maëlle et Arthur, qui ont réussi à me soutenir tout le long, grâce à leur énergie, leur bonne humeur et leurs idées créatives. Ils ont toujours bien voulu m'écouter et me conseiller. Ils m'ont aidé dans les moments les plus durs, réussissant aisément à me motiver.

Travailler avec eux fut vraiment bien.

Et ils sont nombreux au LJK ceux que je voudrais également remercier, ceux que j'ai côtoyés pendant ces trois années et avec qui j'ai partagé de super moments, à la salle à café, au CROUS, au bureau, dans les couloirs, en ville, partout, tout le temps. À eux, merci.





# Table des matières

<b>Introduction générale</b>	<b>9</b>
<b>I Méthodologie et outils</b>	<b>13</b>
<b>1 Assimilation de données</b>	<b>15</b>
1.1 Une brève introduction à l'assimilation de données . . . . .	16
1.1.1 Le principe de l'assimilation de données . . . . .	16
1.1.2 La gestion des erreurs en assimilation de données . . . . .	18
1.2 Différentes méthodes d'assimilation de données . . . . .	19
1.2.1 Le BLUE et l'assimilation de données variationnelle . . . . .	19
1.2.2 L'assimilation de données par filtrage . . . . .	23
1.2.3 Généralisation des formulations variationnelles . . . . .	25
1.3 Assimilation d'images et erreurs de positions . . . . .	26
1.3.1 Les images et données denses en assimilation . . . . .	26
1.3.2 Les erreurs de position . . . . .	29
1.3.3 Gérer les erreurs de position . . . . .	30
<b>2 Transport optimal</b>	<b>33</b>
2.1 Définition du transport optimal . . . . .	34
2.1.1 Densités de probabilité . . . . .	34
2.1.2 Transport entre deux densités de probabilité . . . . .	34
2.1.3 Transport optimal et la distance de Wasserstein . . . . .	35
2.2 Quelques caractéristiques de la distance de Wasserstein . . . . .	37
2.2.1 Une distance sur la position de l'information . . . . .	37
2.2.2 Interpoler à l'aide de la distance de Wasserstein . . . . .	41
2.3 Autres définitions de la distance de Wasserstein . . . . .	44
2.3.1 Le problème de Monge . . . . .	44
2.3.2 Le problème de Kantorovich . . . . .	49
2.3.3 Dualité de Kantorovich . . . . .	50
2.3.4 Résumé . . . . .	53
2.4 L'espace de Wasserstein et son produit scalaire . . . . .	53
2.5 Différentielle de la distance de Wasserstein . . . . .	55
2.6 Résolution numérique du problème de transport optimal . . . . .	57
2.7 Transport optimal pour des densités de masses différentes . . . . .	58
2.7.1 Premières initiatives de transport optimal de densités de masses différentes . . . . .	58
2.7.2 Distance de Wasserstein-Fisher-Rao . . . . .	58

2.7.3	L'espace Riemannien associé à la distance Wasserstein-Fisher-Rao . . . . .	60
2.8	La distance $\mathcal{H}^{-1}$ . . . . .	63
2.8.1	Définition classique de la distance $\mathcal{H}^{-1}$ . . . . .	64
2.8.2	Intérêt : mesurer le déplacement . . . . .	65
2.8.3	Les distances $\mathcal{H}^{-1}$ et lien avec la distance de Wasserstein . . . . .	65
2.8.4	Barycentre pour la distance $\mathcal{H}^{-1}$ . . . . .	67
2.9	Conclusion . . . . .	67
<b>II Transport Optimal pour l'Assimilation de Données</b>		<b>69</b>
<b>3</b>	<b>Assimilation de données à l'aide du transport optimal</b>	<b>71</b>
3.1	État de l'art de l'assimilation de données et des problèmes inverses utilisant le transport optimal . . . . .	72
3.2	Présentation de l'assimilation de données utilisant le transport optimal . . . . .	73
3.2.1	Cas général . . . . .	73
3.2.2	Présentation des expériences qui vont suivre . . . . .	75
3.3	Minimisation d'une fonction coût de Wasserstein . . . . .	76
3.3.1	La méthode de descente de plus forte pente . . . . .	76
3.3.2	Descente du gradient euclidien de $\mathcal{J}(\rho_0)$ . . . . .	78
3.3.3	Descente du gradient de Wasserstein . . . . .	80
3.3.4	Formulation géodésique de la méthode de descente de plus forte pente . . . . .	83
3.3.5	Descente du gradient avec régularisation . . . . .	84
3.3.6	Résumé des méthodes de descente . . . . .	85
3.4	Descente de plus forte pente pour une fonction $\mathcal{H}^{-1}$ . . . . .	87
3.5	Les autres méthodes de descente du gradient . . . . .	88
3.5.1	La méthode du gradient conjugué . . . . .	88
3.5.2	Les méthodes de Newton et de quasi-Newton . . . . .	89
3.5.3	Implémentation numérique . . . . .	89
<b>4</b>	<b>Contrôle d'une densité de probabilité</b>	<b>91</b>
4.1	Choix du terme d'ébauche . . . . .	91
4.2	Description des expériences suivantes . . . . .	94
4.3	Contrôle d'une seule densité dans un modèle linéaire. . . . .	95
4.4	Assimilation dans un modèle de Saint-Venant . . . . .	95
4.5	Ajout d'un terme de régularisation . . . . .	97
4.6	Capacité à retrouver la forme de l'objet . . . . .	99
4.7	Observations soumises à du bruit . . . . .	102
4.8	Dans le cas de présence d'erreur d'amplitude seulement . . . . .	103
4.9	Distance Wasserstein-Fisher-Rao pour des densités positives . . . . .	105
4.10	Conclusion de ces contrôles d'une seule densité de probabilité . . . . .	106
<b>5</b>	<b>Assimilation dans un modèle complet de Saint-Venant</b>	<b>109</b>
5.1	Contrôle de la vitesse initiale . . . . .	111
5.1.1	Contrôle de la vitesse initiale seule et sans terme d'écart à l'ébauche . . . . .	111
5.1.2	Contrôles de la vitesse et de la densité initiales . . . . .	113
5.1.3	En présence d'erreurs de position . . . . .	116
5.2	Contrôle du fond du bassin . . . . .	117

5.2.1	Contrôle du fond seul, sans terme d'ébauche . . . . .	118
5.2.2	Contrôles du fond et de la hauteur de la colonne d'eau avec ébauche . . . . .	119
5.2.3	Utilisation d'une distance euclidienne pour le terme d'ébauche . . . . .	121
5.3	Contrôle de toutes les variables en même temps . . . . .	122
5.4	Conclusion générale de ces expériences . . . . .	125
<b>6</b>	<b>Conclusion et perspectives</b>	<b>127</b>
6.1	Résumé du travail effectué . . . . .	127
6.2	Perspectives . . . . .	131
6.2.1	Produit scalaire alternatif pour le fond . . . . .	131
6.2.2	$n$ -dimension . . . . .	132
6.2.3	Statistiques d'erreurs . . . . .	135
6.2.4	Occultations . . . . .	138
6.2.5	Assimilation de données par filtrage et par la distance de Wasserstein . . . . .	139
<b>A</b>	<b>Résolution numérique du transport optimal</b>	<b>141</b>
A.1	Cas uni-dimensionnel . . . . .	141
A.2	Versions primale, duale et primale-duale du transport optimal . . . . .	142
A.3	Résolution du problème primal par descente de gradient . . . . .	144
A.4	Algorithme pour la version duale . . . . .	146
A.5	Résolution du problème primal-dual . . . . .	147
A.6	Régularisation d'entropie . . . . .	151
A.7	Algorithme pour un calcul numérique de la distance WF. . . . .	154
<b>B</b>	<b>La méthode de l'adjoint</b>	<b>159</b>
B.1	Gradient d'une fonction impliquant un modèle . . . . .	159
B.2	Le modèle tangent . . . . .	160
B.3	Adjoint continu . . . . .	161
B.4	Adjoint discret . . . . .	162
B.5	Équations adjointes pour modèle de Saint-Venant avec fond . . . . .	162
<b>C</b>	<b>Dans le cas où les conditions aux bords ne sont pas de Neumann</b>	<b>165</b>
C.1	Définition du produit scalaire . . . . .	166
	<b>Liste des tableaux</b>	<b>169</b>
	<b>Table des figures</b>	<b>171</b>
	<b>Bibliographie</b>	<b>174</b>



# Introduction générale

Les prévisions météorologiques jouent un rôle majeur dans le quotidien des populations. Leurs intérêts sont multiples. Elles constituent pour certaines professions un véritable repère sur lequel on doit pouvoir s'appuyer. Par exemple, les agriculteurs usent des prévisions météorologiques afin de gérer au mieux l'eau nécessaire à l'arrosage des cultures. Par ailleurs, elles jouent un rôle protecteur envers la population en diffusant des bulletins d'alertes des prochaines intempéries telles que les journées de forte chaleur, les tempêtes, etc.. Par exemple, minimiser les dégâts importants que peuvent causer les cyclones est d'enjeu crucial et est rendu possible grâce aux prévisions.

L'origine des prévisions météorologiques est militaire. Le 14 novembre 1854, pendant la guerre de Crimée, une forte tempête éclata dans la mer Noire et coula 38 navires français. Or d'après un certain Urbain Le Verrier, alors directeur de l'Observatoire de Paris, cette catastrophe aurait pu être évitée : en effet, la tempête avait traversé auparavant toute l'Europe et si différentes observations avaient été relevées et transmises, son arrivée en Crimée aurait pu être anticipée. Ainsi, avec le directeur général des lignes télégraphiques et à la demande expresse de Napoléon III, Urbain Le Verrier fonda le bureau de météorologie à l'Observatoire de Paris. Ce bureau regroupa initialement les observations provenant par télégraphe de 24 stations météorologiques implantées en France, pour s'étendre à 65 stations implantées dans toute l'Europe, grâce auxquelles un bulletin météorologique journalier put être édité.

Ces prévisions météorologiques sont faites à la main et se basent sur les tendances que suivent les observations. En 1922, Lewis Fry Richardson développe un premier modèle numérique d'évolution météorologique. Les équations sont discrétisées et résolues à la main, l'ordinateur n'existant pas encore. Dans ce premier modèle, la justesse des prévisions nécessite la connaissance de la valeur des différentes variables de l'atmosphère au temps initial. Les experts lui fournissent alors des conditions initiales dites "subjectives" après étude visuelle des différentes mesures prélevées. Avec l'arrivée des ordinateurs, quelques algorithmes ont été développés et permettent d'initialiser automatiquement le modèle météorologique, comme l'analyse de Cressman (1950). Avec ces premiers algorithmes débuta l'ère de l'assimilation de données.

L'assimilation de données peut être définie comme l'ensemble des méthodes servant à estimer la condition initiale (et par extension des paramètres incertains) d'un modèle physique à partir de toutes les informations disponibles sur celui-ci.

La qualité des prévisions météorologiques n'a cessé de croître au fil des années. L'amélioration de la description des modèles et l'augmentation de la performance des ordinateurs y ont largement contribué. La qualité et la quantité croissantes des observations faites de l'atmosphère, ainsi que leur assimilation toujours plus efficace grâce au développement de différentes méthodes d'assimilation de données fournissant des conditions initiales de plus en plus précises, sont autant de facteurs expliquant cette évolution.

Les observations météorologiques ont aujourd’hui différentes sources : stations météorologiques, avions de ligne, ballons, qui fournissent des mesures ponctuelles ; mais aussi des satellites fournissant des images de la Terre. Ces images satellites sont les données les plus cruciales pour la prévision, comme le prouve la figure 1 qui fournit la contribution des différentes sources d’observations aux prévisions météorologiques. Les deux principales sources sont les instruments AMSU-A et IASI du satellite européen MetOp, dont les images contribuent à eux seuls à 42% de la prévision.

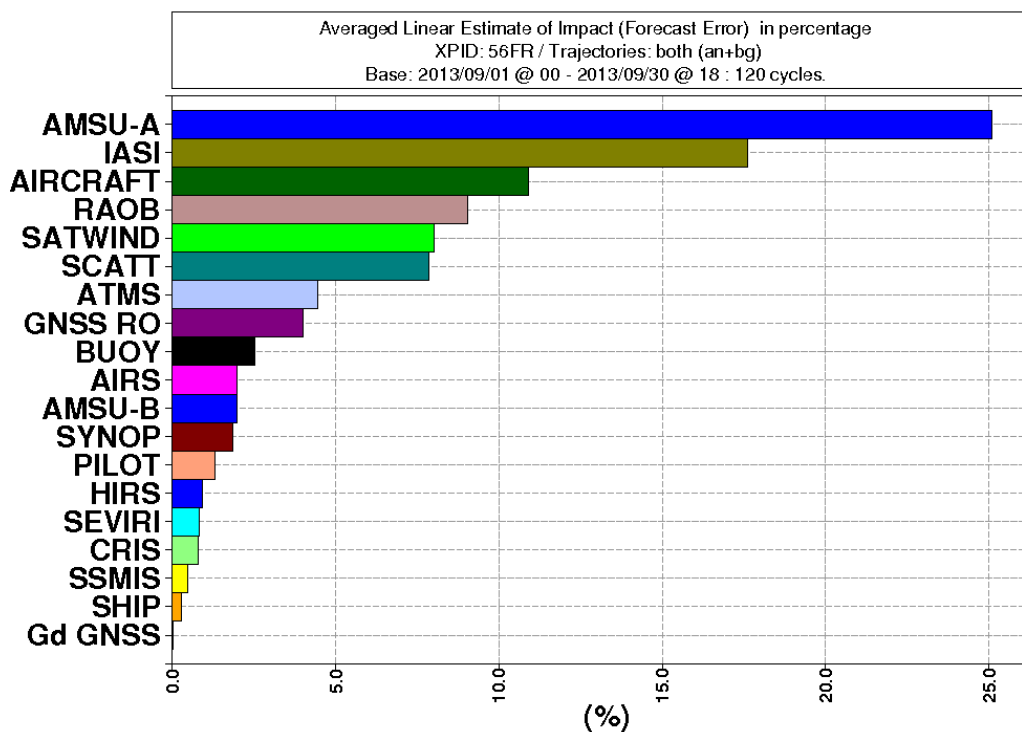


FIGURE 1 – Contributions des différentes sources d’observations à l’analyse en météorologie, pour une prévision de Septembre 2013 à Météo France. On constate que les deux premières sources d’information sont issues d’images satellites, à 42% environ de l’impact sur l’analyse (AMSU-A et IASI sont deux instruments du satellite MetOp). Les autres principales sources d’information sont issues des avions de ligne (AIRCRAFT - 11%), des ballons (RAOB - 9%) et des cartes de vent issues notamment d’AMV (SATWIND - 6%). Copyright (2016) IROWG.

Plus généralement, il faut pour tous les domaines usant de la modélisation mathématique estimer la condition initiale et les paramètres et donc assimiler un certain nombre de données. Ainsi, l’assimilation de données est utilisée dans les domaines de l’océanographie, la climatologie, le médical (simulation du cœur humain), l’imagerie, l’aéronautique, la neutronique (estimation du flux neutronique au cœur d’une centrale nucléaire), et bien d’autres. L’essor de l’imagerie dû aux progrès techniques a rendu l’utilisation de données images fréquente dans ces domaines. Par exemple, il est de plus en plus courant d’utiliser des caméras pour faire des mesures, ou d’avoir recours à des scanners en médecine.

Contrairement aux données ponctuelles, les images ont une cohérence spatiale qui n’est actuellement pas utilisée pour leur assimilation directe. En effet, l’assimilation directe des données passe par leurs comparaisons, comparaisons nécessitant le choix d’une distance. La distance utilisée classiquement pour les images est une distance pixel à pixel et n’utilise pas cette cohérence. Un problème survient notamment dans le cas où les données possèdent des erreurs de position, c’est-à-dire si

les structures des images sont mal placées. C'est une erreur que la distance classique ne peut pas capturer et cela résulte en une condition initiale déformée. Dans le cadre de cette thèse, nous avons testé plusieurs distances prenant en compte la cohérence spatiale en vue d'une meilleure assimilation dans le cas d'erreurs de position. Ces distances ont par exemple la propriété de prendre en compte le positionnement relatif des structures.

Nous nous concentrerons en particulier sur la distance de Wasserstein issue de la théorie du transport optimal, domaine mathématique dont le cœur a été développé il y a une trentaine d'années. Cette distance de Wasserstein a beaucoup d'applications en traitement d'images et possède des propriétés intéressantes d'interpolation. Nous verrons comment cette distance peut aider à assimiler les images.

Nous nous intéresserons dans ce manuscrit à l'assimilation d'images mais aussi à d'autres types de données à cohérence spatiale, à l'aide de ces distances. Les travaux de cette thèse sont nés au sein de l'équipe MOISE puis AIRSEA chez Inria et suivent ceux de [Souopgui, 2010, Chabot, 2014] traitant aussi d'assimilation d'images.

## Organisation du manuscrit

Dans le chapitre 1, nous allons présenter et définir clairement les méthodes d'assimilation de données. Nous parlerons de données à cohérence spatiale qui peuvent présenter des erreurs de position et de leur assimilation. Nous allons notamment insister sur le choix des distances permettant les comparaisons entre les données. Le chapitre 2 sera consacré aux distances calculant l'écart entre deux images. Nous y présenterons la théorie du transport optimal de laquelle naît la distance de Wasserstein. Nous y évoquerons aussi la distance Wasserstein-Fisher-Rao et la distance  $\mathcal{H}^{-1}$ .

Des méthodes d'assimilation utilisant ces distances seront présentées en détail dans le chapitre 3. Les spécificités de ces méthodes ainsi que la comparaison avec la méthode classique seront présentées à travers quelques premiers exemples d'abord dans le chapitre 4, puis à travers d'autres scénarios plus réalistes dans le chapitre 5.

Enfin, le chapitre 6 de conclusion discutera de l'applicabilité de cette méthode à l'assimilation de données réelles.





Première partie

Méthodologie et outils



# Chapitre 1

## Assimilation de données

### Contenu du chapitre

1.1	Une brève introduction à l’assimilation de données . . . . .	16
1.1.1	Le principe de l’assimilation de données . . . . .	16
1.1.2	La gestion des erreurs en assimilation de données . . . . .	18
1.2	Différentes méthodes d’assimilation de données . . . . .	19
1.2.1	Le BLUE et l’assimilation de données variationnelle . . . . .	19
1.2.2	L’assimilation de données par filtrage . . . . .	23
1.2.3	Généralisation des formulations variationnelles . . . . .	25
1.3	Assimilation d’images et erreurs de positions . . . . .	26
1.3.1	Les images et données denses en assimilation . . . . .	26
1.3.2	Les erreurs de position . . . . .	29
1.3.3	Gérer les erreurs de position . . . . .	30

Nous nous intéressons à un système physique, dont on décrit mathématiquement le comportement temporel. Pour le représenter au temps  $t$ , nous disposons d’un ensemble de variables spécifiques mathématiques, ensemble que l’on appelle “état”. Nous disposons de plus d’un modèle mathématique représentant son évolution, formé par des équations différentielles faisant varier en temps la valeur de l’état.

Pour connaître l’état du système au temps  $t > 0$ , il est nécessaire de connaître l’état au temps  $t = 0$  (la condition initiale au modèle mathématique). Nous pourrions penser à utiliser des instruments mesurant sa valeur au temps  $t = 0$ ; cependant, à cause de la taille du système, ceci n’est généralement pas possible. Alors, il faudra combiner plusieurs mesures, mais celles-ci sont peut-être encore incomplètes, anisochrones, bruitées, contradictoires... Reconstituer directement l’état du système à partir de l’ensemble des mesures existantes n’est pas facile et il faut avoir recours au domaine mathématique de l’assimilation de données.

L’assimilation de données est une branche des problèmes inverses dédiée à l’estimation de l’état d’un système physique, ou des quantités importantes pour sa description, en combinant toutes les informations que l’on a sur ce système. Sa volonté première était de fournir des conditions initiales à des modèles météorologiques pour des prévisions. Nous étudierons les enjeux de l’assimilation dans la section 1.1 et quelques méthodes dans la section 1.2.

Les observations des systèmes physiques sont de plus en plus de type image. Les systèmes d’imagerie observant les systèmes physiques se sont développés et sont bien utilisés maintenant. Par exemple, les images provenant des satellites sont de plus en plus nombreuses, ou le recours à des scanners en imagerie médicale (tomographie) est maintenant fréquent.

Nous verrons dans la section 1.3 comment sont assimilées ces images aujourd’hui. Nous verrons que pour une assimilation directe, les erreurs de position dans les données peuvent impliquer des mauvaises analyses. Pour résoudre ce problème, nous proposerons d’utiliser d’autres distances.

Les éléments de ce chapitre se sont appuyés sur les différents travaux de thèse suivants : [Souopgui, 2010, Bonan, 2013, Chabot, 2014, Mirouze, 2010, Papadakis, 2007, Rochoux et al., 2014a].

## 1.1 Une brève introduction à l’assimilation de données

### 1.1.1 Le principe de l’assimilation de données

Les notations d’assimilation de données utilisées ici suivent les recommandations classiques de [Ide et al., 1997].

On dispose d’un système physique (par exemple l’atmosphère) décrit en tout temps  $t$  et en tout point de l’espace  $x$  par un ensemble de variables, appelé **état** et qu’on note  $\mathbf{x}(t, x)$ , avec  $t$  le temps et  $x$  la position. Ce vecteur d’état  $\mathbf{x}(t, x)$  contient notamment l’ensemble des variables nécessaires pour décrire le système. Pour l’atmosphère, on peut penser aux variables d’humidité, de pression, de température, et la vitesse du vent, entre autres. L’état peut aussi inclure les conditions aux bords lorsque le domaine est borné, ou la valeur de certaines constantes (les paramètres) qui interviennent dans sa description mathématique. La valeur de  $\mathbf{x}$  est inconnue puisqu’on ne dispose pas d’instruments pouvant la mesurer dans sa totalité et sans erreurs. L’estimer est important, pour connaître son état futur (la prévision, comme en météorologie), son état présent, ou même son état passé. Le dernier cas est de la réanalyse : par exemple la climatologie cherche à déterminer l’évolution passée de la composition et de la température de l’atmosphère, comme dans [Uppala et al., 2005] où l’évolution du climat de ces 40 dernières années est estimée.

L’évolution spatio-temporelle du système physique est décrite par un modèle mathématique que l’on note  $\mathcal{M}$ . Généralement, il s’écrit sous forme d’équations différentielles :

$$\mathcal{M}: \begin{cases} \frac{\partial \mathbf{x}}{\partial t} = \mathcal{F}(t, \mathbf{x}) \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases} \quad (1.1.1)$$

où  $\mathcal{F}$  est une fonction du temps et de l’état, et  $\mathbf{x}_0$  est l’état initial du système.

Ainsi, connaître l’état au temps  $t > 0$  est équivalent à le connaître au temps  $t = 0$  : nous nous contenterons ainsi de retrouver  $\mathbf{x}_0$ , la condition initiale du modèle, qui donnera  $\mathbf{x}(t, x)$  par simple intégration du modèle. Nous écrivons que  $\mathbf{x} = \mathcal{M}(\mathbf{x}_0)$ .

#### Remarque 1.1.1

*Cependant, il y a souvent un écart entre le système physique et le modèle mathématique le décrivant, étant donné que ce modèle est une simplification de la réalité. De plus, le modèle mathématique doit être discrétisé pour pouvoir être résolu numériquement, ce qui ajoute une potentielle source d’erreur. Par exemple, le modèle météorologique actuel est discrétisé sur une maille dont le pas est d’ordre kilométrique, faute de mieux, ce qui ne permet pas de représenter les phénomènes locaux de plus petite échelle qui peuvent avoir une influence sur le comportement global.*

*Nous reviendrons sur cette difficulté dans la prochaine section 1.1.2.*

Pour pouvoir estimer  $\mathbf{x}_0$ , on utilise des observations  $\mathbf{y}^{\text{obs}} = (\mathbf{y}_i^{\text{obs}})_{i=1..N_{\text{obs}}}$  qui sont des mesures du système prises à différents temps  $t_1, \dots, t_{N_{\text{obs}}}$ . Ces observations ne mesurent pas toutes les variables du système à la fois; et elles mesurent parfois une variable qui n'appartient pas au vecteur d'état. Par exemple en océanographie on observe souvent la concentration de phytoplancton qui n'intervient pas dans le modèle. Les mesures sont effectuées soit en un point du système (mesure locale *via* une bouée ou un bateau en mer par exemple), soit dans une zone du système (comme une image satellitaire de la surface de l'océan), soit sur tout le domaine dans lequel vit le système.

Retrouver  $\mathbf{x}_0$  fait appel à des méthodes d'assimilation de données.

**Définition 1.1.2 (*L'assimilation de données*)**

*L'assimilation de données est l'ensemble des méthodes et algorithmes permettant d'estimer l'état  $\mathbf{x}_0$  d'un système, à l'aide d'un modèle mathématique décrivant l'évolution du système, des observations disponibles, des statistiques décrivant les erreurs, et de toute autre information possible sur cet état.*

Les techniques d'assimilation vont chercher à s'approcher de la vraie valeur de l'état que l'on note  $\mathbf{x}_0^t$  ( $t$  pour *true* en anglais). Elles vont faire en sorte que  $\mathbf{x} := \mathcal{M}(\mathbf{x}_0)$  soit proche des observations  $\mathbf{y}^{\text{obs}}$ . Comme  $\mathbf{x}$  et  $\mathbf{y}^{\text{obs}}$  ne vivent pas dans le même espace, il est nécessaire d'avoir recours à un opérateur d'observation  $\mathcal{H}$  qui va calculer des équivalents-observations  $\mathbf{y}$  à partir de  $\mathbf{x}$ ,

$$\mathbf{y} := \mathcal{H}(\mathcal{M}(\mathbf{x}_0)).$$

Pour simplifier les notations, nous introduisons l'opérateur général d'observations  $\mathcal{G}$  défini par

$$\mathcal{G}(\mathbf{x}_0) := \mathcal{H}(\mathcal{M}(\mathbf{x}_0)).$$

Une première approche naïve de l'assimilation de données serait de trouver la condition initiale passant par les observations, c'est-à-dire le  $\mathbf{x}_0$  tel que

$$\mathcal{G}(\mathbf{x}_0) = \mathbf{y}^{\text{obs}}. \tag{1.1.2}$$

Ce problème inverse souffre de quelques défauts :

- Les observations ne sont pas parfaites. Elles peuvent être entachées de bruit à cause des imperfections de l'instrument de mesure, ou parce qu'elles ont été créées par combinaison d'autres observations (par exemple dans le cas de *pseudo-observations*, nous y reviendrons), et l'étape de combinaison peut être la résolution inexacte d'un problème inverse. Ainsi, si les observations sont imparfaites, on peut ne pas vouloir résoudre exactement 1.1.2.
- Ce problème est mal posé, puisque *sous-déterminé*. Le nombre de valeurs à estimer (l'inconnue  $\mathbf{x}_0$ ) est souvent bien plus important que le nombre de valeurs observées. Par exemple en météorologie, après discrétisation, on veut estimer  $10^7$  variables en ayant uniquement  $10^6$  valeurs d'observations. Une conséquence est que le nombre de solutions peut être infini.

Par conséquent, la solution au problème (1.1.2) est mal définie. Pour pallier ce problème, il est commun d'ajouter de l'information, en supposant que la solution  $\mathbf{x}_0$  doit être également proche d'une ébauche  $\mathbf{x}_0^b$  (avec  $b$  pour *background* en anglais). Cette ébauche a deux utilités :

- La première utilité est mathématique et est de sur-déterminer le problème. L'ébauche a un rôle de régularisation de Tikhonov en problèmes inverses. Le problème devient alors de trouver un état initial  $\mathbf{x}_0$  étant proche de l'ébauche et passant proche des observations. Comme nous le verrons, cela s'écrit en un problème bien posé.

- La deuxième utilité est physique et permet d'incorporer de l'information supplémentaire dans le problème. Il est par exemple courant d'avoir déjà une connaissance *a priori* de l'état. Si l'on sait que la solution ressemble à telle autre solution on peut le dire à travers cette ébauche. Par exemple si l'on a déjà fait une prévision de l'état auparavant, cette prévision peut servir comme ébauche.

Pour résumer, le problème devient de trouver un  $\mathbf{x}_0$  qui serait un compromis entre les informations apportées par les observations d'une part et par l'ébauche d'autre part. On voudrait formuler le problème comme ceci :

$$\text{trouver } \mathbf{x}_0 \text{ tel que : } \quad \begin{array}{ll} \mathbf{x}_0 & \text{soit proche de } \mathbf{x}_0^b \\ \mathcal{G}(\mathbf{x}_0) & \text{soit proche de } \mathbf{y}^{\text{obs}}. \end{array} \quad (1.1.3)$$

Ce problème sera mieux défini après avoir compris le rôle des erreurs dans la prochaine section, et sera résolu par une méthode d'assimilation de données. Le résultat de la méthode utilisée sera nommé "analyse", notée  $\mathbf{x}_0^a$ .

Résumons les différentes notations impliquées dans le problème :

- l'état vrai du système  $\mathbf{x}^t$  représentant la réalité qu'on ne peut pas complètement mesurer et que l'on cherche à retrouver ;
- les observations  $\mathbf{y}^{\text{obs}} = (\mathbf{y}_i^{\text{obs}})_{i=1..N_{\text{obs}}}$  sont une mesure de l'état vrai et serviront à l'estimer ;
- le modèle  $\mathcal{M}$  simulant l'évolution du système, modèle qu'on a supposé parfait et qui permet de décrire la trajectoire  $\mathbf{x}^t$  uniquement à l'aide de la connaissance de  $\mathbf{x}_0^t$  ;
- l'opérateur d'observation  $\mathcal{H}$  qui calcule l'équivalent observation d'un état, pour passer de l'espace des états à l'espace des observations ;
- l'opérateur général d'observations  $\mathcal{G} := \mathcal{H} \circ \mathcal{M}$  ;
- l'ébauche  $\mathbf{x}_0^b$ , élément en général utile pour que le problème inverse soit bien posé ;
- enfin, l'analyse  $\mathbf{x}_0^a$  est l'estimateur de l'état vrai et est résultat de la méthode d'assimilation de données utilisée.

### 1.1.2 La gestion des erreurs en assimilation de données

Nous avons défini l'assimilation de données comme la résolution d'un problème de compromis entre les observations et les ébauches. Pour définir un bon compromis entre ces données (entachées d'erreur), on peut s'intéresser à leur précision. L'ébauche est une estimation inexacte de la vraie solution et comporte des erreurs. De même, les observations ainsi que le modèle sont incertains.

L'hypothèse générale d'assimilation de données est de considérer ces erreurs comme gaussiennes et de moyenne nulle. De plus les erreurs d'observation et d'ébauche sont supposées indépendantes. Ainsi nous écrivons que les données  $\mathbf{x}_0^b$  et  $\mathbf{y}^{\text{obs}}$  s'expriment en fonction de  $\mathbf{x}_0^t$  par :

$$\begin{array}{lll} \mathbf{y}^{\text{obs}} & = & \mathcal{G}(\mathbf{x}_0^t) + \varepsilon^{\text{obs}} \quad \text{avec } \varepsilon^{\text{obs}} \sim \mathcal{N}(0, \mathbf{R}) \\ \mathbf{x}_0^b & = & \mathbf{x}_0^t + \varepsilon^b \quad \text{avec } \varepsilon^b \sim \mathcal{N}(0, \mathbf{B}). \end{array} \quad (1.1.4)$$

Les matrices  $\mathbf{R}$  et  $\mathbf{B}$  sont les matrices de covariance d'erreur respectivement de  $\varepsilon^{\text{obs}}$  et de  $\varepsilon^b$ . Les erreurs  $\varepsilon^{\text{obs}}$  et  $\varepsilon^b$  sont indépendantes.

Le modèle, comme on l'a dit, n'est qu'une approximation du processus physique décrit. Le décalage dû à cette approximation est l'erreur modèle, défini précisément comme la différence entre  $\mathbf{x}^t(t + \Delta t)$  et  $\mathcal{M}_{t \rightarrow t + \Delta t}(\mathbf{x}^t(t))$  (l'intégration du modèle entre  $t$  et  $t + \Delta t$ ). Il est généralement supposé que cette erreur suit une loi normale d'espérance nulle et de matrice de covariance  $\mathbf{Q}_t$ .

Dans le reste du manuscrit nous négligerons ces erreurs modèle et nous poserons  $\mathbf{Q}_t = 0$ .

## 1.2 Différentes méthodes d'assimilation de données

Maintenant que nous avons introduit le problème, voyons plus en détail quelques méthodes d'assimilation de données qui vont déterminer l'analyse. Il y a les méthodes variationnelles (cf. la section 1.2.1) et les méthodes de filtrage (cf. la section 1.2.2). Le choix des méthodes et des approximations est large comme nous allons le voir en 1.2.3.

Nous rappelons que le but de l'assimilation de données est de trouver une condition initiale  $\mathbf{x}_0$  au modèle décrivant le système physique à étudier, en accord avec les différentes informations disponibles. Ce problème peut être vu notamment comme trouver un compromis entre d'une part les observations et d'autre part l'ébauche, tout en suivant les équations du modèle.

### 1.2.1 Le BLUE et l'assimilation de données variationnelle

Une première version classique est de considérer  $\mathbf{x}_0$  comme une correction de  $\mathbf{x}_0^b$ . Cette correction peut être linéaire et s'écrit

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{K}(\mathbf{y}^{\text{obs}} - \mathcal{G}(\mathbf{x}_0^b))$$

avec  $\mathbf{K}$  une matrice de gain à déterminer : c'est le principe du BLUE (*Best Linear Unbiased Estimator*), sur lequel les autres méthodes d'assimilation de données se baseront, comme nous allons le voir après. Le BLUE suppose que l'opérateur  $\mathcal{G}$  est linéaire (on l'écrit alors  $\mathbf{G}$  pour suivre les notations de [Ide et al., 1997]). Dans ce cas, la correction à faire est estimée en fonction des matrices de covariance d'erreur et de  $\mathbf{G}$ . Par exemple si l'erreur sur les observations est importante (une grande covariance  $\mathbf{R}$ ), mais que l'ébauche est assez précise (faible covariance  $\mathbf{B}$ ), on fera plus confiance en l'ébauche.

Le choix de  $\mathbf{K}$  est déterminé en minimisant l'erreur de  $\mathbf{x}_0$  (la trace de sa matrice de covariance d'erreur). Nous passons les détails que l'on peut retrouver par exemple dans [Blayo et al., 2016], qui conduisent au BLUE :

#### Méthode 1.2.1 (*Le BLUE (Best Linear Unbiased Estimator)*)

La méthode du BLUE consiste à écrire

$$\mathbf{x}_0 = \mathbf{x}_0^b + \mathbf{K}(\mathbf{y}^{\text{obs}} - \mathbf{G}\mathbf{x}_0^b)$$

avec la matrice de gain  $\mathbf{K}$  valant

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1}. \quad (1.2.1)$$

C'est une méthode simple puisqu'il n'y a que la matrice  $\mathbf{K}$  à calculer (matrice qui peut cependant être de trop grande taille). Cette méthode n'est par contre pas définie pour les modèles  $\mathcal{G}$  non-linéaires. Le BLUE est en fait généralisable en constatant qu'il est équivalent dans le cas linéaire au 4D-Var (décrit ci-dessous) ainsi qu'au problème de maximum *a posteriori* qui est :

$$\underset{\mathbf{x}_0}{\text{maximiser}} P^a(\mathbf{x}_0). \quad (1.2.2)$$



avec  $P^a(\mathbf{x}_0)$  la densité de probabilité d'être proche de  $\mathbf{x}_0^t$  en connaissance (*a posteriori*) des observations. Ce problème conduit au 4D-Var :

### Méthode 1.2.2 (*Méthode variationnelle du 4D-Var*)

Le problème de maximum *a posteriori* (1.2.2) est équivalent au problème variationnel suivant, nommé 4D-Var :

$$\underset{\mathbf{x}_0}{\text{minimiser}} \mathcal{J}(\mathbf{x}_0) := \|\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}}\|_{\mathbf{R}}^2 + \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}}^2 \quad (1.2.3)$$

avec les normes  $\|\cdot\|_{\mathbf{R}}$  et  $\|\cdot\|_{\mathbf{B}}$  des normes euclidiennes pondérées par les matrices  $\mathbf{R}^{-1}$  et  $\mathbf{B}^{-1}$ , respectivement définies pour  $\mathbf{x}$  et  $\mathbf{y}$  par

$$\|\mathbf{y}\|_{\mathbf{R}}^2 = \mathbf{y}^T \mathbf{R}^{-1} \mathbf{y}, \quad \|\mathbf{x}\|_{\mathbf{B}}^2 = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x}.$$

De plus, ces deux problèmes sont équivalents au BLUE dans le cas où  $\mathcal{G}$  est linéaire.

### Démonstration des équivalences :

**Équivalence entre le 4D-Var et le problème de maximum *a posteriori*.** Ce résultat repose sur le théorème de Bayes. En effet, on écrit :

$$P^a(\mathbf{x}_0) = P(\mathbf{x}_0 = \mathbf{x}_0^t | \mathbf{y} = \mathbf{y}^{\text{obs}}) = \frac{P(\mathbf{y} = \mathbf{y}^{\text{obs}} | \mathbf{x}_0 = \mathbf{x}_0^t) P(\mathbf{x}_0 = \mathbf{x}_0^t)}{P(\mathbf{y} = \mathbf{y}^{\text{obs}})}. \quad (1.2.4)$$

Le dénominateur  $P(\mathbf{y} = \mathbf{y}^{\text{obs}})$  ne dépend pas de  $\mathbf{x}_0$ , il peut être oublié comme nous maximisons sur  $\mathbf{x}_0$ . Le terme  $P(\mathbf{y} = \mathbf{y}^{\text{obs}} | \mathbf{x}_0 = \mathbf{x}_0^t)$  est calculable, puisque dans le cas où l'état vaut  $\mathbf{x}_0$ , les observations valent  $\mathcal{G}(\mathbf{x}_0)$  et

$$P(\mathbf{y} = \mathbf{y}^{\text{obs}} | \mathbf{x}_0) = P(\mathcal{G}(\mathbf{x}_0) = \mathbf{y}^{\text{obs}}).$$

C'est la densité de probabilité que  $\mathcal{G}(\mathbf{x}_0)$  vaille  $\mathbf{y}^{\text{obs}}$  qui est donnée analytiquement par :

$$P(\mathcal{G}(\mathbf{x}_0) = \mathbf{y}^{\text{obs}}) = C \exp(-(\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}})^T \mathbf{R}^{-1} (\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}})) = C \exp(-\|\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}}\|_{\mathbf{R}}^2)$$

avec  $C$  une constante positive dépendant uniquement de la dimension et du déterminant de  $\mathbf{R}$ . De même,  $P(\mathbf{x}_0 = \mathbf{x}_0^t)$ , l'estimateur *a priori*, est donné à l'aide de l'ébauche,

$$P(\mathbf{x}_0 = \mathbf{x}_0^t) = C' \exp(-(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b)) = C' \exp(-\|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}}^2)$$

avec  $C'$  positif et indépendant de  $\mathbf{x}_0$ . Le problème (1.2.2) est équivalent à minimiser l'opposé du logarithme de (1.2.4), et en enlevant les constantes qui ne dépendent pas de  $\mathbf{x}_0$ , conduit à (1.2.3).

**Équivalence entre le 4D-Var et BLUE dans le cas où le modèle est linéaire.** On se place dans le cas où  $\mathcal{G}$  est linéaire et où il s'écrit  $\mathbf{G}$ . Le gradient de  $\mathcal{J}$  est nul au minimum, gradient qui s'écrit :

$$\frac{1}{2} \text{grad} \mathcal{J}(\mathbf{x}_0) = \mathbf{G}^T \mathbf{R}^{-1} (\mathbf{G} \mathbf{x}_0 - \mathbf{y}) + \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b)$$

Soit  $\mathbf{x}_0$  atteignant le minimum de  $\mathcal{J}$ . Comme le gradient y est nul, on a :

$$\underbrace{(\mathbf{G}^T \mathbf{R}^{-1} \mathbf{G} + \mathbf{B}^{-1})}_{\mathbf{A}} \mathbf{x}_0 = \mathbf{G}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{B}^{-1} \mathbf{x}_0^b.$$

Donc en inversant  $\mathbf{A}$ , et en remarquant que  $\mathbf{A}^{-1} \mathbf{B}^{-1} = \mathbf{I} - \mathbf{A}^{-1} \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G}$ , on a :

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{A}^{-1} \mathbf{G}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{x}_0^b - \mathbf{A}^{-1} \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G} \mathbf{x}_0^b \\ &= \mathbf{x}_0^b + \mathbf{A}^{-1} \mathbf{G}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{G} \mathbf{x}_0^b) \\ &= \mathbf{x}_0^b + \mathbf{K} (\mathbf{y} - \mathbf{G} \mathbf{x}_0^b) \end{aligned}$$

avec le  $\mathbf{K}$  explicité dans le BLUE, dans (1.2.1). Ceci prouve l'équivalence entre le 4D-Var et le BLUE dans le cas où  $\mathcal{G}$  est linéaire.

Ce résultat est la base de l'assimilation variationnelle de données : pour trouver un "bon" état, on minimise la fonction coût donnée dans (1.2.3). Cela consiste à trouver l'élément le plus proche possible au sens de la métrique associée aux erreurs, à la fois des observations et de l'ébauche, voir la figure 1.1. Cette méthode s'appelle le 4D-Var et a été formulée premièrement dans les années 50 dans [Sasaki et al., 1955, Sasaki, 1958]. On dit "4D" pour parler des trois dimensions spatiales et de la dimension temporelle mises en jeu. Les premiers algorithmes numériques de résolution pour la météorologie ont été décrits dans [Le Dimet et Talagrand, 1986] et ont été opérationnels dès 1993 à l'ECMWF [Courtier et al., 1998].

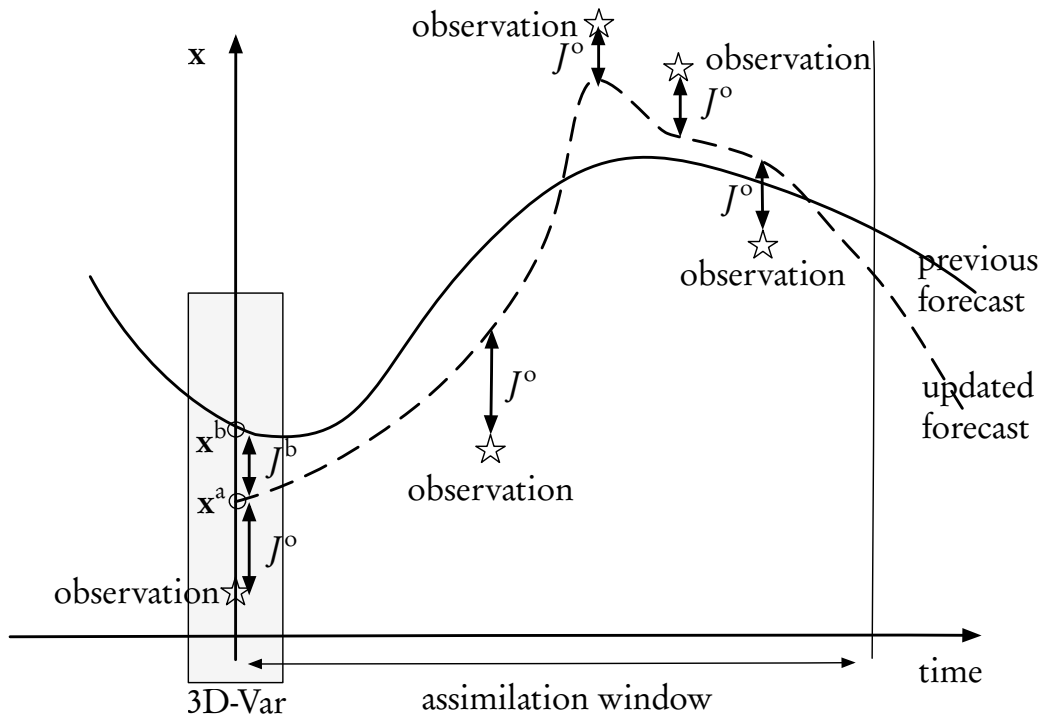


FIGURE 1.1 – Le principe du 4D-Var est de trouver une analyse  $\mathbf{x}^a$  donnant la courbe temporelle en pointillé, qui soit à la fois proche de l'ébauche  $\mathbf{x}^b$  dont l'évolution temporelle est en trait plein, et des observations représentées par des étoiles.

Le 4D-Var s'apparente à un problème de moindres carrés qui est très commun en problèmes inverses. On peut remarquer que le terme  $\|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}}^2$  joue le rôle d'une régularisation de Tikhonov.

Regardons de plus près les mises en œuvres numériques associées à  $\mathcal{J}$  et sa minimisation.

### Remarque 1.2.3 (Implications numériques sur les matrices $\mathbf{R}$ et $\mathbf{B}$ )

Pour calculer  $\mathcal{J}$ , il faut avoir accès à  $\mathbf{R}^{-1}$  et  $\mathbf{B}^{-1}$ . Dans la plupart des cas, la taille de ces matrices est la même que celle des données, donc les inverser est très coûteux voire impossible. Ainsi, des simplifications sont nécessaires.

- La matrice  $\mathbf{B}$  est généralement remplacée par un opérateur qui ne calcule que  $\mathbf{B}\mathbf{x}$ . Il n'y a donc pas de problèmes de stockage. Par exemple,  $\mathbf{B}$  est un opérateur de diffusion. Voir

[Mirouze, 2010] pour plus de détails.

- La matrice  $\mathbf{R}$  est généralement considérée comme diagonale, pour pouvoir facilement l'inverser. Une matrice  $\mathbf{R}$  diagonale représente des données d'observations complètement décorrélées : la covariance entre deux observations est nulle mais aussi la covariance entre deux points différents d'une même mesure est nulle. Cette hypothèse est très forte car ce n'est pas le cas dans la réalité.

#### Remarque 1.2.4 (La minimisation numérique de $\mathcal{J}$ )

La minimisation (numérique) de  $\mathcal{J}$  est un point crucial qu'il est nécessaire d'explicitier. Cela repose sur des méthodes de minimisation basées sur le gradient que nous détaillerons pour notre problème dans le chapitre 3. Le calcul du gradient de  $\mathcal{J}$  est rendu efficace grâce à la méthode de l'adjoint [Le Dimet et Talagrand, 1986], décrite aussi dans l'Annexe B. Essentiellement, cela conduit à résoudre une équation différentielle rétrograde et de la même taille que celle du modèle  $\mathcal{M}$ . Les méthodes de minimisation de  $\mathcal{J}$  basées sur le gradient sont du type

$$\mathbf{x}_0^{n+1} = \mathbf{x}_0^n + \mathbf{d}^n$$

avec  $\mathbf{d}^n$  la direction de descente dépendant de la valeur du gradient de  $\mathcal{J}$ . Avec un bon choix de  $\mathbf{d}^n$ , la suite  $\mathcal{J}(\mathbf{x}_0^n)$  décroît jusqu'à la valeur minimale et  $\mathbf{x}_0^{n+1}$  tend vers le minimiseur et ce de façon plus ou moins rapide selon l'algorithme utilisé.

Si  $\mathcal{J}(\mathbf{x}_0)$  est (fortement) strictement convexe, la convergence de ces algorithmes vers le minimum global est assurée. En pratique c'est rarement le cas à cause des non-linéarités du modèle et l'algorithme de minimisation peut tomber dans un minimum local. Il existe des solutions, comme la méthode du 4D-Var multi-incrémental décrite ci-dessous avec lequel on pourra quand même finir dans un minimum local mais proche du minimum global.

Du 4D-Var tel qu'il est présenté dans la méthode 1.2.2 découle plusieurs méthodes, dont certaines sont vues comme des simplifications.

**Le 4D-Var incrémental.** Cette méthode proposée dans [Courtier et al., 1994] repose sur une linéarisation autour de  $\mathbf{x}_0^b$  de  $\mathcal{G}$ , pour n'avoir à trouver qu'un incrément  $\delta\mathbf{x}_0$ . En effet, on a :

$$\mathcal{G}(\mathbf{x}_0^b + \delta\mathbf{x}_0) = \mathcal{G}(\mathbf{x}_0^b) + \mathbf{G}[\mathbf{x}_0^b].\delta\mathbf{x}_0 + \mathcal{O}(\delta\mathbf{x}_0)$$

et l'on suppose que  $\delta\mathbf{x}_0$  est petit. C'est la méthode de Gauss-Newton, où l'on transforme ainsi l'opérateur non-linéaire  $\mathcal{G}$  en un opérateur linéaire, ce qui transforme le problème en :

$$\mathcal{J}(\delta\mathbf{x}_0) = \|\mathbf{G}[\mathbf{x}_0^b].\delta\mathbf{x}_0 - \mathbf{d}\|_{\mathbf{R}}^2 + \|\delta\mathbf{x}_0\|_{\mathbf{B}}^2 \quad (1.2.5)$$

avec  $\mathbf{d} = \mathbf{y}^{\text{obs}} - \mathcal{G}(\mathbf{x}_0^b)$  le vecteur d'innovation. Cette fonction coût est convexe et peut être minimisée efficacement. Une fois que le  $\delta\mathbf{x}_0$  optimal est trouvé, on a une analyse  $\mathbf{x}_0^a = \mathbf{x}_0^b + \delta\mathbf{x}_0$ . On peut réitérer ce procédé en linéarisant autour de ce  $\mathbf{x}_0^a$  et trouver un autre  $\delta\mathbf{x}_0$ , etc.. **Le 4D-Var multi-incrémental** proposé dans [Veersé et Thepaut, 1998] simplifie le problème puisqu'il propose de résoudre itérativement  $\delta\mathbf{x}_0$  sur des grilles de plus en plus fines, ou sur de plus en plus de modes propres. Cette procédure empêche aussi de tomber dans des minima locaux trop loin de la solution.

Ces simplifications font que le 4D-Var incrémental est beaucoup utilisé en météorologie, par exemple pour ECMWF, MétéoFrance ou pour l'initialisation du modèle global de Metoffice<sup>1</sup> (couplé avec une méthode séquentielle).

1. Voir <http://www.metoffice.gov.uk/research/modelling-systems/unified-model/weather-forecasting>

**Le 3DVar.** Le 4D-Var est simplifié en posant simplement que le modèle  $\mathcal{M}$  est l'identité, supprimant ainsi la dimension temporelle. Cela donne un **3DVar** plus simple à résoudre, que l'on peut encore simplifier en linéarisant l'opérateur  $\mathcal{H}$ , donnant du **3DVar incrémental** ou du **3DVar multi-incrémental**. On peut se permettre cette simplification lorsque la fenêtre d'assimilation est courte, lorsque le modèle n'influe pas trop sur la condition initiale; par exemple dans le modèle AROME de MétéoFrance ou le modèle UKV<sup>1</sup> de Metoffice, qui sont des modèles de prévision à quelques heures sur une échelle fine (kilométrique). On retrouve le 3DVar pour des modèles d'océanographie également.

Nous avons présenté quelques méthodes d'assimilation variationnelle de données qui résident en la minimisation d'une fonction coût. Cette fonction coût dépend directement des erreurs sur les données. Avant de voir qu'il existe un large choix de fonction coût, nous allons présenter l'autre type de méthodes d'assimilation de données, l'assimilation par filtrage qui a été beaucoup développée ces dernières années.

### 1.2.2 L'assimilation de données par filtrage

L'autre catégorie de méthodes d'assimilation de données, regroupant les méthodes par filtrage, est brièvement répertoriée ici. Les observations  $(\mathbf{y}_i^{\text{obs}})_{i=1..N^{\text{obs}}}$  sont prises aux temps  $(t_i)_{i=1..N^{\text{obs}}}$ . On écrit que le modèle  $\mathcal{M}$  qui intègre sur toute la fenêtre d'assimilation, peut être décomposé en petits éléments  $\mathcal{M}_i$ , chacun intégrant sur la fenêtre  $[t_i, t_{i+1}]$ . De même, l'état  $\mathbf{x}$  est décomposé en  $\mathbf{x}_i = \mathbf{x}(t_i)$ . Donc en partant d'une condition initiale  $\mathbf{x}_0$ , l'état aux temps  $t_i$  est calculé simplement par l'algorithme récursif

$$\forall 0 \leq i \leq N^{\text{obs}} - 1, \quad \mathbf{x}_{i+1} = \mathcal{M}_i(\mathbf{x}_i).$$

Le filtre de Kalman [Kalman, 1960] puis le filtre de Kalman-Bucy [Kalman et Bucy, 1961] ont été développés vers 1960, et ont par exemple servi à guider avec succès la fusée Apollo. Ils sont très utilisés aujourd'hui. Ils permettent à chaque temps  $t_i$  de corriger la valeur de  $\mathbf{x}_i$  en utilisant les informations disponibles (telles que l'observation  $\mathbf{y}_i^{\text{obs}}$ ), voir la figure 1.2. Cette correction se base exactement sur le BLUE. L'algorithme du filtre de Kalman s'écrit :

#### Méthode 1.2.5 (*Filtre de Kalman*)

*Le filtre de Kalman est composé de 3 étapes :*

**Prévision.** *L'état au temps  $t_{i+1}$  est prédit, sachant l'état au temps  $i$  et les  $i$  premières observations,*

$$\hat{\mathbf{x}}_{i+1|i} = \mathcal{M}_i(\hat{\mathbf{x}}_i).$$

**Correction.** *Cet état au temps  $t_{i+1}$  ne tient pas compte de l'observation  $i + 1$ . Pour en tenir compte, on effectue un BLUE, ou 3DVar, écrit comme*

$$\hat{\mathbf{x}}_{i+1} = \arg \min_{\mathbf{x}_{i+1}} \left[ \|\mathcal{H}_{i+1}(\mathbf{x}_{i+1}) - \mathbf{y}_{i+1}^{\text{obs}}\|_{\mathbf{R}_{i+1}}^2 + \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}_{i+1|i}\|_{\mathbf{B}_{i+1}}^2 \right]. \quad (1.2.6)$$

**Calcul des matrices de covariance d'erreur.** *La matrice  $\mathbf{R}_i$  est connue à chaque  $i$  puisqu'elle est donnée avec les observations. Par contre, la matrice  $\mathbf{B}_i$  est inconnue, c'est la matrice de l'erreur sur  $\hat{\mathbf{x}}_{i+1|i}$ . Dans le cas de modèles et opérateurs d'observations linéaires,*

elle est calculée par récurrence à l'aide du BLUE via les trois matrices suivantes :

$$\begin{aligned} \text{Matrice de gain associée à } \hat{\mathbf{x}}_i : \quad & \mathbf{K}_i = (\mathbf{B}_i^{-1} + \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i)^{-1} \mathbf{H}_i^T \mathbf{R}_i^{-1} \\ \text{Covariance d'erreur de } \hat{\mathbf{x}}_i : \quad & \mathbf{P}_i = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{B}_i \\ \text{Covariance d'erreur de } \hat{\mathbf{x}}_{i+1|i} : \quad & \mathbf{B}_{i+1} = \mathbf{M}_i \mathbf{P}_i \mathbf{M}_i^t. \end{aligned}$$

Dans le cas de modèles non-linéaires, la matrice  $\mathbf{B}_i$  est estimée par linéarisation ou en utilisant des méthodes dites ensemblistes (voir le paragraphe ci-dessous).

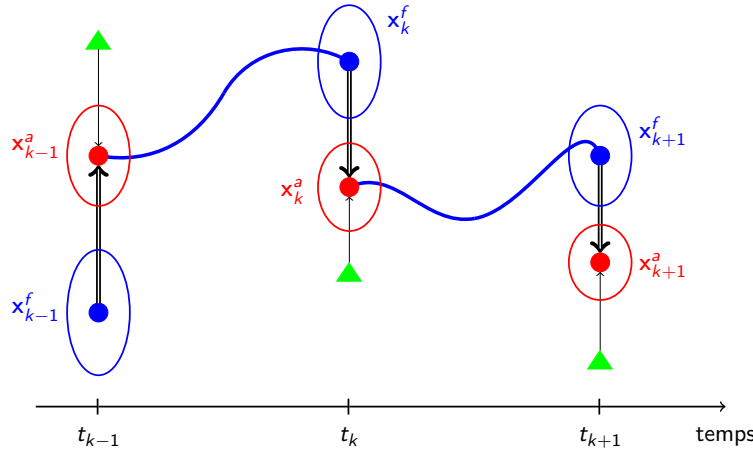


FIGURE 1.2 – Représentation schématique du filtre de Kalman : l'analyse  $\mathbf{x}_{k-1}^a$  au temps  $k-1$  donne la prédiction  $\mathbf{x}_k^f$  au temps  $k$  à l'aide du modèle, qui est corrigée à l'aide de l'observation (le triangle vert), donnant l'analyse  $\mathbf{x}_k^a$ , ... Ainsi de suite jusqu'à la dernière observation.

Les deux méthodes (le 4D-Var et le filtre de Kalman) donnent la même analyse à la fin de la fenêtre temporelle si le modèle est linéaire et parfait (pas d'erreur modèle), voir [Li et Navon, 2001]. Si le modèle est imparfait (si le modèle ne décrit pas parfaitement la réalité), on modélise l'erreur entre la prédiction et l'intégration du modèle par une loi gaussienne de matrice de covariance  $\mathbf{Q}_i$ , qui doit être incluse dans le calcul de  $\mathbf{B}_i$ .

**Le filtre de Kalman d'ensemble (EnKF).** Le problème du filtre de Kalman réside dans le calcul de la matrice  $\mathbf{B}_i$ , qui est au temps  $t_i$  une simple propagation par le modèle des erreurs représentées par  $\mathbf{P}_{i-1}$  au temps  $t_{i-1}$ . Cela nécessite au modèle d'être linéaire, cas auquel on peut se ramener en linéarisant le modèle. Mais même dans le cas linéaire, il s'avère coûteux de calculer  $\mathbf{B}_i$  à cause de sa taille. En effet, il faut intégrer le modèle autant de fois que la dimension des matrices  $\mathbf{P}_i$ .

Pour contourner ce problème, on utilise un ensemble de particules qui permet d'approcher cette matrice, à l'aide d'une estimation stochastique. Le procédé est comme suit : d'abord, des particules sont générées par un échantillonnage de type Monte-Carlo. Ensuite chacune de ces particules est intégrée par le modèle direct, jusqu'au temps d'observation suivant. Enfin, la valeur de chaque particule est corrigée à l'aide d'un BLUE, offrant la matrice de covariance entre les particules qui approche la matrice de covariance souhaitée.

Ces méthodes par filtrage seront exploitées dans le chapitre 6 et dans une bien moindre mesure que les méthodes variationnelles tout au long du manuscrit. Nous allons voir maintenant que les méthodes variationnelles sont très diverses, puisque les fonctions coût à minimiser possibles sont nombreuses. Nous verrons notamment l'application de l'assimilation d'images.

### 1.2.3 Généralisation des formulations variationnelles

Dans le cadre de cette thèse nous nous focaliserons sur le 4D-Var décrit dans la partie 1.2.1. La fonction coût qui est utilisée classiquement en assimilation variationnelle de données est

$$\mathcal{J}(\mathbf{x}_0) = \|\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}}\|_{\mathbf{R}}^2 + \|\mathbf{x}_0 - \mathbf{x}_0^b\|_{\mathbf{B}}^2 \quad (1.2.7)$$

avec  $\mathbf{R}$  et  $\mathbf{B}$  les matrices de covariance d'erreur. Minimiser cette fonction coût exprime le fait qu'on veut trouver un état étant un compromis entre les observations et l'ébauche. On peut en fait écrire plus généralement le compromis de deux ensembles de données comme la solution du problème

$$\underset{\mathbf{x}_0}{\text{minimiser}} \mathcal{J}(\mathbf{x}_0) := d_o(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + d_b(\mathbf{x}_0, \mathbf{x}_0^b)^2 \quad (1.2.8)$$

où  $d_o$  et  $d_b$  sont deux distances à préciser, respectivement dans l'espace des observations et l'espace d'état. Dans (1.2.7) les distances utilisées sont des distances dites de Mahalanobis, des distances euclidiennes pondérées par les matrices  $\mathbf{B}^{-1}$  et  $\mathbf{R}^{-1}$ .

Quel que soit le choix de  $d_o$  et  $d_b$ , donc de la métrique qu'on utilise, le minimum sera toujours un élément "au milieu" de l'ébauche et des observations. Il y a plusieurs raisons et avantages à utiliser de nouvelles distances.

- Cela permet d'avoir une fonction convexe plus simple à minimiser. Dans certains cas, utiliser une distance euclidienne fait apparaître un grand nombre de minima locaux. Le problème de sismologie étudié dans [Engquist et Froese, 2013, Métivier et al., 2016] montre l'avantage que l'on a à changer de distance.
- On peut représenter les données dans d'autres espaces parfois plus adaptés. Il est possible d'écrire les erreurs sur un changement de variable des données avant d'écrire la fonction coût. En effet si on appelle  $\mathcal{T}$  une transformation sur les observations, on peut avoir

$$\mathcal{T}(\mathbf{y}^{\text{obs}}) = \mathcal{T}(\mathcal{H}(\mathbf{x}_0^t)) + \tilde{\varepsilon}^{\text{obs}} \quad \text{avec} \quad \tilde{\varepsilon}^{\text{obs}} \sim \mathcal{N}(0, \tilde{\mathbf{R}})$$

avec  $\tilde{\mathbf{R}}$  une matrice de covariance d'erreur, diagonale, et qui n'est pas forcément liée à  $\mathbf{R}$ . Cela implique un terme dans la fonction coût qui sera

$$\mathcal{J}_o(\mathbf{x}_0) = \|\mathcal{T}(\mathbf{y}^{\text{obs}}) - \mathcal{T}(\mathcal{H}(\mathbf{x}_0))\|_{\tilde{\mathbf{R}}}^2$$

et donc une autre définition de la distance sur les données.

Par exemple dans [Chabot, 2014], on applique aux variables de contrôle un gradient ou une transformation en ondelettes ou en curvelettes pour pouvoir écrire la fonction coût à l'aide des distances respectives à ces espaces. Dans ces espaces on se rend compte que la matrice de covariance d'erreur est à diagonale dominante, donc manipulable.

Dans [Rochoux et al., 2014b] les auteurs proposent une autre distance que la distance classique. Les auteurs ont besoin de mesurer la distance entre deux fronts de feu, représentés dans une image comme dans la figure 1.3. Classiquement on utiliserait la norme de la différence entre deux images de feu, mais étant donné que le front est simplement une courbe, la distance ne varierait pas en fonction de la position des fronts. Les auteurs utilisent plutôt une distance euclidienne locale mesurant l'écart, le déplacement, entre les fronts de feu pour permettre une correction anisotrope.

- Nous verrons dans la suite de ce manuscrit que changer de distance permet aussi d’avoir des minima de formes plus intéressantes. En effet, le minimum de  $\mathcal{J}$  est fonction des choix des distances  $d_o$  et  $d_b$ . Nous présenterons plusieurs cas où un mauvais choix de ces distances amène à un minimum, qui est certes optimal pour  $\mathcal{J}$ , mais qui sera de forme peu désirée. Nous étudierons ce point particulier dans le chapitre 4 notamment.

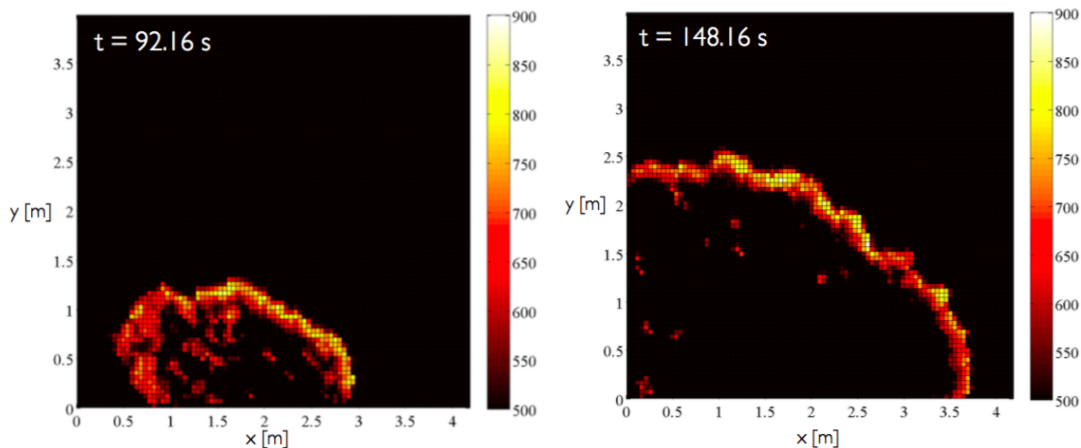


FIGURE 1.3 – Un feu se propageant dans un pré, ayant commencé quelque part en bas. On a les images à  $t = 92,16s$  et  $t = 148,16s$  de la température du pré et donc la localisation du feu. On observe notamment un front de feu, une ligne, qui se déplace vers les zones non encore incendiées. Source : [Rochoux et al., 2014a].

Le principal problème de l’assimilation variationnelle de données est l’écriture de la fonction coût, *i.e.* le choix des distances. Un exemple crucial dans le cadre de ce manuscrit est le cas des images. Une de leurs spécificités est leur cohérence spatiale et nous verrons une distance capable de la prendre en compte.

## 1.3 Assimilation d’images et erreurs de positions

### 1.3.1 Les images et données denses en assimilation

Nous nous intéressons dans cette partie à l’assimilation d’images en particulier, ou de données denses en espace en général.

#### Définition 1.3.1 (*Donnée dense*)

*Une donnée dense (en espace) est une donnée (ébauche ou observation) ayant suffisamment de valeurs pour avoir une cohérence spatiale. On la traitera dans ce manuscrit comme une donnée continue en espace, et que l’on notera dans la suite  $\rho(x)$ . Même si les données denses ne sont pas toutes des images, nous dirons dans la suite du manuscrit par simplicité d’écriture que les données denses sont des images.*

Voir la figure 1.4 pour un exemple de donnée dense. On peut penser aussi à la valeur de la hauteur d’eau dans l’océan qui n’est pas une image mais qu’on pourra considérer comme une donnée dense.

Les images ont de spécifique qu’elles sont bidimensionnelles et qu’il ne manque pas de données (pas de trous).

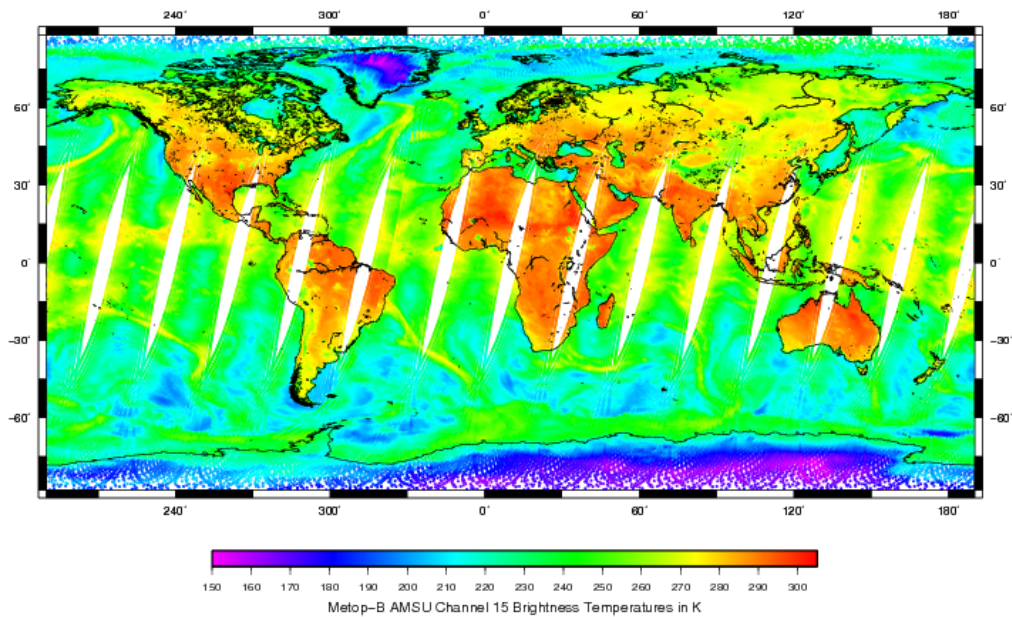


FIGURE 1.4 – Une image satellite de la température de la surface de la Terre, prise par l’instrument AMSU-A du satellite MetOp. C’est une donnée dense, avec des trous, et recomposée puisque le satellite ne capture pas la Terre d’un seul coup. Copyright (2016) EUMETSAT.

On dira que l’ébauche  $\mathbf{x}_0^b$  et de surcroît  $\mathbf{x}_0$  dépendent de la variable spatiale  $x \in \Omega$ , le domaine spatial du système étudié. On dispose également d’observations  $\mathbf{y}_i^{\text{obs}}(x)$  qui dépendent de l’espace, avec  $x$  dans  $\Omega_i^{\text{obs}}$ , un espace convexe, compact, et qui peut être différent de  $\Omega$ .

Dans le plus grand domaine d’application de l’assimilation de données, la météorologie, les images sont prépondérantes grâce aux nombreux satellites en orbite fournissant un grand nombre de données de différentes sortes. Ces données sont d’une très grande utilité depuis 1960 avec le lancement de TIROS-1. Aujourd’hui, le service européen EUMETSAT dispose du satellite MetOp, qui est équipé d’une dizaine d’instruments mesurant différentes variables de la Terre.

À noter qu’un satellite ne renvoie pas directement une image. Souvent il faut traiter les données du satellite (comme débruiter, gérer les occlusions dues aux nuages, décoder les différentes fréquences) et cela donne déjà lieu à un problème d’inversion où des erreurs importantes peuvent être générées.

Les images peuvent être représentées de différentes manières. Citons-en deux, qui ne sont pas exclusives, comme nous allons le voir dans la suite du manuscrit.

**Représentation en pixels.** On dit souvent que c’est la représentation élémentaire de l’image. À chaque point de l’image (pixel) on peut attribuer un niveau de gris ou une valeur de rouge, vert et bleu (niveau RGB pour les images de couleur). Cette interprétation brute de l’image ne permet pas d’en tirer des caractéristiques intéressantes. Par exemple, on peut penser que l’information se situe dans les discontinuités. Pour avoir accès à plus d’informations il faut changer de représentation de l’image.



**Décomposition en base d'ondelettes (ou de curvelettes).** En stockant l'image non pas en tant que somme de pixels mais en tant que somme d'ondelettes, on traduit peut-être mieux l'information. Les ondelettes ont été introduites en 1984 [Grossmann et Morlet, 1984] et sont extrêmement utilisées pour représenter l'information d'une image, celle-ci se retrouvant dans les premiers coefficients de la décomposition en ondelette. L'assimilation d'images représentées dans cette base a été étudiée notamment dans [Titaud et al., 2010, Vidard et al., 2014] et offre des résultats intéressants notamment en cas d'occultations, puisque le résultat garde une cohérence spatiale due à cette représentation.

Une conséquence du choix de la représentation de l'image est le choix de la distance utilisée pour les comparer. Par exemple, comparer deux images représentées par leurs pixels utilise la distance euclidienne (la norme de la différence), ou la représentation en ondelettes utilise une distance sur les coefficients d'ondelettes (qui se trouve être la même que la distance euclidienne). Dans [Titaud et al., 2010] les auteurs utilisent une distance sur une troncature de la base de curvelettes, offrant une simplification de la distance et du calcul de  $\mathcal{J}$ .

Nous nous intéresserons dans ce manuscrit à d'autres choix de distances spécifiques aux images.

Enfin, les images sont assimilées en pratique de deux manières différentes. Soit elles sont directement assimilées comme de simples observations, soit elles sont transformées et regroupées en ce qu'on appelle des pseudos-observations. On va voir leurs influences pour l'assimilation de données en météorologie.

**Les pseudos-observations** Une première méthode d'assimilation d'images est d'extraire l'information dynamique d'une séquence d'observations, et d'assimiler uniquement cette information dynamique. Cela a l'avantage de réduire le nombre d'images à assimiler en plus d'avoir une information assez fiable puisqu'elle regroupe plusieurs observations. Les cartes de vent obtenues s'appellent les AMV<sup>2</sup>. Par exemple, à partir d'une dynamique d'images de couvertures nuageuses (cf. la figure 1.5), l'idée est d'en extraire le champ de vitesse.

La méthode classique pour obtenir l'information dynamique est la PIV<sup>3</sup> [Adrian, 1991]. C'est une méthode statistique couramment utilisée en physique expérimentale pour obtenir le champ de vitesse.

Une autre méthode est le flot-optique. Introduite en 1981 dans [Horn et Schunck, 1981], cette méthode est beaucoup utilisée depuis 1993 [Schmetz et al., 1993] en météorologie ou en océanographie, voir quelques exemples dans [Huot et al., 2006, Auroux et Fehrenbach, 2011].

La construction de l'AMV peut aussi utiliser d'autres mesures, comme les données synoptiques des stations météorologiques ou des mesures satellites de la vitesse des vents de surface réalisées par le diffusomètre ASCAT du satellite MetOp.

De nombreux algorithmes ont été développés pour avoir une carte des vents avec le moins d'erreur et le moins de corrélation entre les erreurs possibles [Dew et Holmlund, 2000, Deb et al., 2008, Kishtawal et al., 2009]. En effet, un problème de ces cartes est la grande corrélation spatiale et temporelle de l'erreur [Bormann et al., 2003] qui les rend délicates à utiliser.

**L'assimilation directe des images.** L'apport des satellites en image est important. L'instrument le plus utile à l'assimilation en météorologie est l'AMSU-A qui mesure la température de

---

2. AMV pour *Atmospheric Motion Vector*

3. PIV pour *Particle Image Velocities*

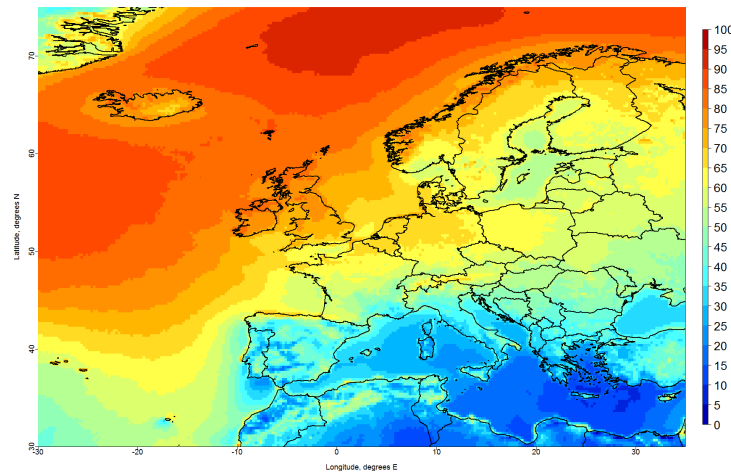


FIGURE 1.5 – Image de la couverture nuageuse (exprimée en pourcentage) du ciel, mesurée par l'instrument AVHRR du satellite MetOp. Copyright (2016) EUMETSAT.

la surface de la Terre avec une grande précision, voir la figure 1.4. On note aussi l'IASI qui mesure les profils de la composition de l'air, de la température et de l'humidité dans l'atmosphère. Ces deux instruments fournissent des images qui sont directement assimilées, sans transformation [Zhao et al., 2005, Hilton et al., 2012].

Vu la quantité et la taille de ces données, il est d'usage de les réduire, par exemple en les moyennant (*superobbing* en anglais) ce qui a aussi pour effet de réduire l'erreur d'observation, ou en les sous-échantillonnant (en réduisant la résolution des images, *thinning* en anglais).

### 1.3.2 Les erreurs de position

Une caractéristique des images et données denses est qu'elles ont une cohérence spatiale. Ainsi on compte deux types d'erreurs : l'erreur d'amplitude et l'erreur de position. Sur une image, on peut distinguer plusieurs structures qui, comparées à la réalité, pourront avoir des erreurs d'amplitude (pas la bonne valeur) et/ou des erreurs de position (pas la bonne place).

Dans la gestion classique des données, on suppose qu'elles n'ont que des erreurs d'amplitude, voir par exemple la formulation des erreurs (1.1.4) : on a supposé qu'un bruit additif gaussien a été ajouté aux données, ce qui traduit bien le fait que l'erreur est considérée comme étant uniquement d'amplitude. Or comme on l'a dit, les images sont aussi sujettes à des erreurs de position, qui sont alors traitées simplement comme des erreurs d'amplitude.

Nous tâcherons dans cette partie de déterminer l'impact de ce choix de gestion des erreurs : quelles sont les origines des erreurs de position ? Qu'implique une erreur de position sur l'analyse, lorsque ces erreurs sont traitées comme des erreurs d'amplitude ? Quelles possibilités existe-t-il pour traiter directement les erreurs de position ?

**Les origines des erreurs de position.** Il peut y avoir des erreurs de position à la fois sur les observations, sur l'ébauche et même à cause du modèle.

Les modèles numériques utilisés peuvent conduire à des erreurs de position. Par exemple, si la résolution spatiale ou temporelle est trop faible, un retard ou un décalage dû à cette résolution

se crée et les structures peuvent être mal placées. Le schéma numérique utilisé peut introduire un retard de phase. Des interactions entre les échelles d'un modèle multi-échelle peuvent aussi être l'origine de problème de phase [Ravela et al., 2007].

L'ébauche provient souvent d'un calcul précédent. Mais dans ce calcul précédent des erreurs de position peuvent avoir été introduites : par exemple, si la vitesse d'une structure a été mal estimée, cette structure ne sera pas au bon endroit [Hoffman et Grassotti, 1996]. Des mauvaises estimations constantes aussi, par exemple les constantes barotropiques dans le modèle des ondes de Rossby créent des erreurs de position dans la prévision du Gulf Stream [Mariano, 1990].

Les observations de type images aussi sont sujettes à des problèmes de positions. Cela peut être dû à des problèmes de dérive des instruments de mesure (orientation des instruments satellites). Ou si ces observations sont les résultats d'un calcul (pseudo-observation ou une construction par une étape de *pre-processing*), on peut trouver des déphasages.

### 1.3.3 Gérer les erreurs de position

Différentes études testent l'introduction d'erreurs de position dans les données et montrent que ce type d'erreur peut affecter la qualité de l'analyse bien plus que les erreurs d'amplitude, par exemple dans [Chang et Holt, 1994] ou [Manobianco et al., 1994] pour la prévision de la position d'un cyclone qui se révèle être cruciale pour se préparer au mieux à ce type de catastrophe.

Des tests académiques menés dans [Ravela et al., 2007] montrent que si l'ébauche contient des erreurs de position, alors l'analyse obtenue à l'aide de méthodes classiques (4D-Var ou EnKF) comporte elle des problèmes d'amplitude.

Une explication à cela est que la distance utilisée (la distance euclidienne) ne considère les erreurs de position que comme des erreurs d'amplitude, puisqu'elle mesure la valeur de la différence en amplitude et non en position entre deux images. On peut le voir sur la figure 1.6 où l'on a un exemple de distance euclidienne constante entre des fonctions représentant des objets de positions pourtant bien différentes. Une conséquence est que l'analyse est proche de la moyenne arithmétique des données assimilées. Ainsi, l'analyse va posséder des erreurs d'amplitude, voir la figure 1.7 où l'on voit que la moyenne de données sujettes à des erreurs de position n'est pas aussi ample que les données elles-mêmes.

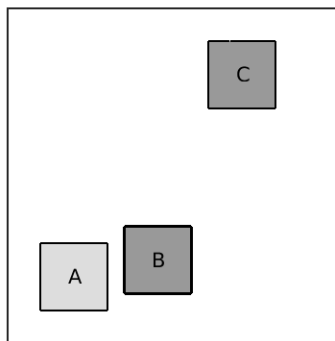


FIGURE 1.6 – Les carrés A, B et C peuvent être représentés par une densité valant 1 dans le carré et 0 ailleurs. La distance euclidienne (la norme de la différence) entre A et B est exactement la même que la distance entre A et C.

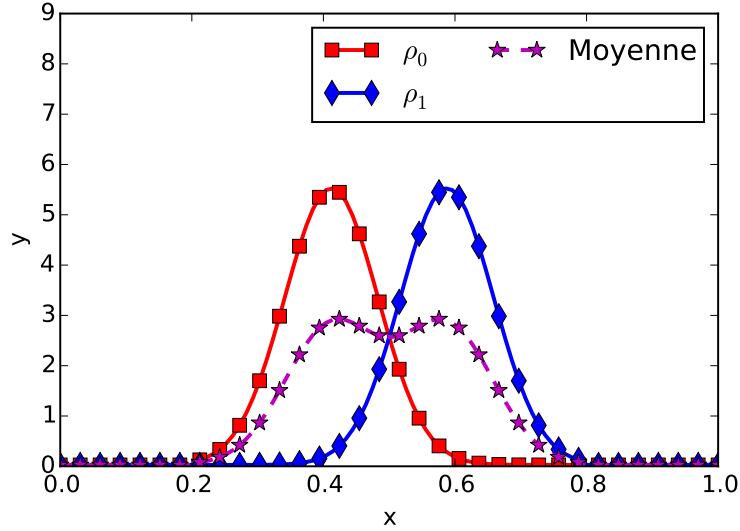


FIGURE 1.7 – Moyenne pour la distance euclidienne de deux densités  $\rho_0$  et  $\rho_1$  sujettes à des erreurs de position. On constate notamment des défauts d'amplitude sur la moyenne, comparée aux données d'origine.

Plusieurs méthodes ont été proposées pour assimiler des données contenant des erreurs de position. Celles-ci consistent généralement à replacer les données pour annuler cette erreur. Dans [Hoffman et Grassotti, 1996] une étape précédant l'assimilation consiste à déplacer l'ébauche d'un vecteur de déplacement  $\mathbf{q}$  (constant en espace) de façon à ce que l'ébauche "colle" avec les observations. Cette idée est approfondie dans [Ravela et al., 2007], où une variable  $\mathbf{q}(x)$  à contrôler est ajoutée à la fonction coût. Cette variable  $\mathbf{q}(x)$  cherche à réaligner l'ébauche. Ainsi, le nouveau coût s'écrit

$$\mathcal{J}(\mathbf{x}_0, \mathbf{q}) = \|\mathcal{G}(\mathbf{x}_0) - \mathbf{y}^{\text{obs}}\|_{\mathbf{R}}^2 + \|\mathbf{x}_0 - \mathbf{x}_0^b(\mathbf{q})\|_{\mathbf{B}(\mathbf{q})}^2 + \mathcal{J}_q(\mathbf{q}).$$

avec  $\mathbf{x}_0^b(\mathbf{q})$  l'ébauche déplacée du vecteur  $\mathbf{q}$  et  $\mathbf{B}(\mathbf{q})$  la matrice de covariance d'erreur sur l'ébauche qui dépend de  $\mathbf{q}$  mais de manière inconnue, d'après les auteurs. Pour l'estimer ils utilisent une méthode d'ensemble. La fonction de pénalisation  $\mathcal{J}_q$  évite au déplacement de l'analyse  $\mathbf{q}^a$  d'être trop grand.

Dans l'idée du choix de la représentation d'une image discutée en section 1.3.1, et du choix de la distance liée, nous proposons d'utiliser de nouvelles distances qui gèrent intrinsèquement les erreurs de position. Ainsi nous écrirons une assimilation variationnelle de données à l'aide de la fonction coût

$$\mathcal{J}(\mathbf{x}_0) := d_o(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + d_b(\mathbf{x}_0, \mathbf{x}_0^b)^2$$

avec  $\mathbf{y}^{\text{obs}}$  des images et  $d_o$ ,  $d_b$  les distances à choisir. Ces distances auront la propriété de calculer l'écart en position entre deux structures. On peut nommer la distance de Wasserstein, issue de la théorie du transport optimal et souvent utilisée en imagerie. Elle a la propriété de déceler les erreurs de position, et pour cette distance, le barycentre de deux structures mal placées est une autre structure placée au milieu, voir la figure 1.8.

La fonction coût à écrire sera du type

$$\mathcal{J}(\mathbf{x}_0) := \mathcal{W}_2(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + \mathcal{W}_2(\mathbf{x}_0, \mathbf{x}_0^b)^2$$

La distance  $\mathcal{H}^{-1}$  a la propriété de mesurer le déplacement entre deux structures mal placées et méritera d'être mentionnée. Elle est liée à la distance de Wasserstein. Ces deux distances seront présentées dans le chapitre suivant.

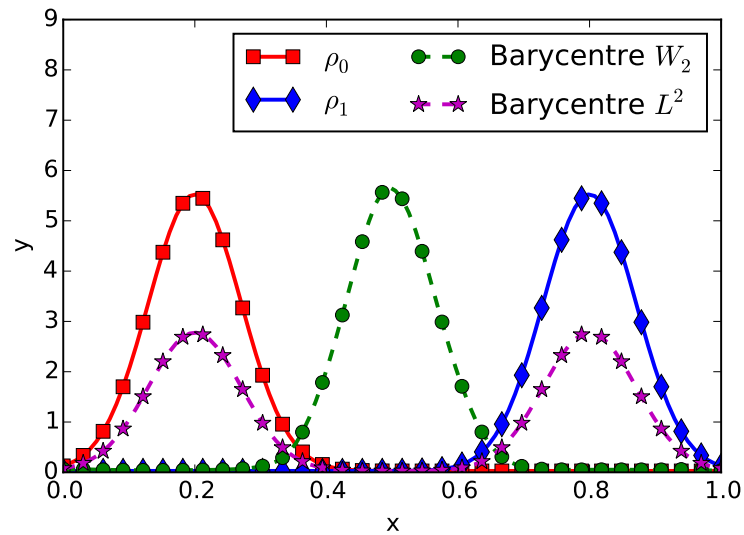


FIGURE 1.8 – Barycentre pour la distance euclidienne  $L^2$  et de Wasserstein  $W_2$  de deux densités  $\rho_0$  et  $\rho_1$  sujettes à des erreurs de position. Le barycentre pour la distance de Wasserstein est une moyenne en position des données d'origine.

# Chapitre 2

## Transport optimal

### Contenu du chapitre

2.1	Définition du transport optimal . . . . .	34
2.2	Quelques caractéristiques de la distance de Wasserstein . . . . .	37
2.3	Autres définitions de la distance de Wasserstein . . . . .	44
2.4	L'espace de Wasserstein et son produit scalaire . . . . .	53
2.5	Différentielle de la distance de Wasserstein . . . . .	55
2.6	Résolution numérique du problème de transport optimal . . . . .	57
2.7	Transport optimal pour des densités de masses différentes . . . . .	58
2.8	La distance $\mathcal{H}^{-1}$ . . . . .	63
2.9	Conclusion . . . . .	67

Nous avons conclu le chapitre précédent en présentant une assimilation de données écrite à l'aide d'une fonction coût de la forme

$$\mathcal{J}(\mathbf{x}_0) = \mathcal{W}_2(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + \mathcal{W}_2(\mathbf{x}_0, \mathbf{x}_0^b)^2$$

en prétendant que cette fonction coût pouvait donner une analyse convenable lorsque les observations sont des images et qu'il y a des erreurs de position. La distance utilisée,  $\mathcal{W}_2$ , est la distance de Wasserstein, issue de la théorie du transport optimal. Le but de ce chapitre est de définir la distance de Wasserstein, de montrer le comportement de  $\mathcal{J}$ , d'exposer les outils permettant de calculer et minimiser  $\mathcal{J}$ . Nous verrons aussi deux distances qui lui sont liées, la distance Wasserstein-Fisher-Rao et la distance  $\mathcal{H}^{-1}$ .

Le transport optimal est défini comme ceci : on dispose de deux fonctions et l'on cherche à transporter l'une sur l'autre de manière à minimiser un coût.

Ce problème a été introduit en 1781 par G. Monge. À cause de la non-linéarité des contraintes de transport, ce problème est longtemps resté sans solution. En 1942, L. Kantorovich pose un problème de transport, plus simple à résoudre, mais équivalent au problème de Monge. La résolution atteint un tournant en 1991, avec la caractérisation de la solution du transport optimal dans un théorème de Yann Brenier (cf. théorème 2.3.1). Dans son sillon, un certain nombre de résultats théoriques sont apparus.

Les applications utilisant le transport optimal sont nombreuses, mais il a fallu attendre que ce problème puisse être calculé numériquement et de manière efficace. Les algorithmes de calcul du

transport optimal furent très longs, jusque dans les années 2000 où les avancées algorithmiques et informatiques ont permis de pallier cet handicap et ainsi de favoriser les applications l'utilisant.

Nous présentons les bases du transport optimal dont nous avons besoin pour l'assimilation de données. Afin de compléter cette introduction, le lecteur pourra se rediriger vers les œuvres [Santambrogio, 2015], ou encore celles de [Villani, 2003, Villani, 2008], plus théoriques.

## 2.1 Définition du transport optimal

Avant de définir le transport optimal, nous définissons dans la section 2.1.1 le type de fonctions transportées : les densités de probabilité. Ensuite, en 2.1.2, nous définirons le transport, et en 2.1.3 nous définirons et caractériserons le transport optimal et la distance de Wasserstein.

La version du transport optimal que l'on expose en premier est la version moderne, celle de [Benamou et Brenier, 2000]. Les versions d'origine, de Monge par exemple, seront présentées plus tard, dans la section 2.3.

### 2.1.1 Densités de probabilité

On définit ici les densités de probabilité qui seront les fonctions à transporter.

Soit  $\Omega \subset \mathbb{R}^d$  ( $d \geq 1$  est la dimension), convexe, compact, de mesure non nulle, et de frontière  $\partial\Omega$  Lipschitzienne. On considère l'espace des mesures de probabilité définies sur  $\Omega$ , c'est-à-dire les mesures positives d'intégrale égale à 1. Pour des raisons numériques, on se cantonne aux mesures de probabilités à densité continue (absolument continues par rapport à la mesure de Lebesgue), c'est-à-dire celles représentées par des fonctions continues. Ainsi, on définit

$$\mathcal{P}(\Omega) = \left\{ \rho \in \mathcal{L}^2(\Omega) : \rho \geq 0 \text{ et } \int_{\Omega} \rho(x) dx = 1 \right\}$$

qui est souvent noté  $\mathcal{P}^{\text{ac}}(\Omega)$  dans la littérature, "ac" pour "absolument continu".

#### Remarque 2.1.1

*En météorologie, on peut représenter la couverture nuageuse terrestre par un élément de  $\mathcal{P}(\Omega)$  (normalisée pour que l'intégrale vaille un), voir la Figure 2.1. En océanographie, le niveau d'eau ou la concentration de chlorophylle peuvent être représentés par une densité de probabilité. Plus généralement, toute image en niveau de gris, où chaque pixel est codé par une valeur entre 0 et 1 peut également être représentée par une fonction de  $\mathcal{P}(\Omega)$ , à condition qu'elle soit normalisée pour que l'intégrale vaille 1.*

### 2.1.2 Transport entre deux densités de probabilité

#### Définition 2.1.2

*Transporter une densité  $\rho_0 \in \mathcal{P}(\Omega)$  sur une autre  $\rho_1 \in \mathcal{P}(\Omega)$  correspond à trouver :*

- *un chemin  $\rho$  de  $\mathcal{P}(\Omega)$  (i.e. tel que  $\forall t \in [0, 1], \rho(t, \cdot) \in \mathcal{P}(\Omega)$ ), continu en temps et en espace, et tel que  $\rho(0) = \rho_0$  et  $\rho(1) = \rho_1$  ;*
- *un champ de vitesse  $v(t, x)$  défini sur  $[0, 1] \times \Omega$ , tel que pour tout  $t$ ,  $v(t, x) \in \mathcal{L}^2(\rho(t))$  ;*

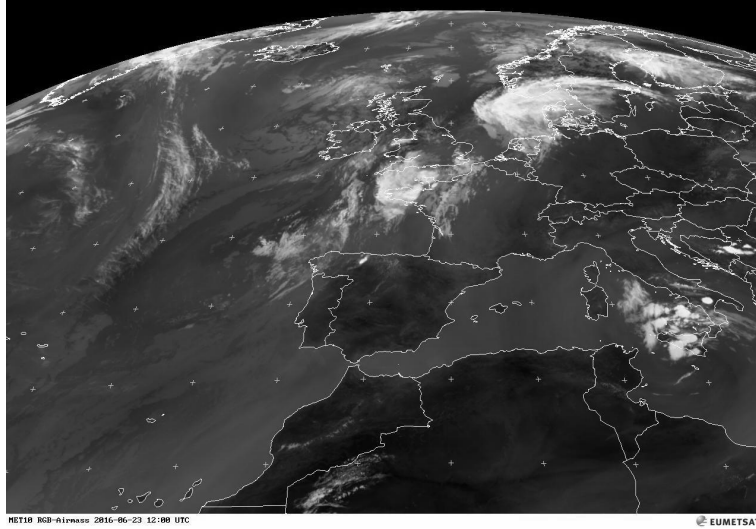


FIGURE 2.1 – La couverture nuageuse se déplaçant sur la surface de la Terre peut en tout instant être représentée par une fonction de  $\mathcal{P}(\Omega)$ . Cette fonction valant 0 là où il n’y a pas de nuage. (Copyright 2016 Eumetsat).

vérifiant l’équation de continuité, ou équation de transport :

$$\begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ \rho(t=0) = \rho_0, \quad \rho(t=1) = \rho_1 \end{cases} \quad (2.1.1)$$

avec les flux aux bords non sortants

$$\rho v \cdot \vec{n} = 0 \text{ presque partout sur } [0, 1] \times \partial\Omega$$

qu’il faut prendre dans sa formulation faible : pour tout  $\Phi \in \mathcal{C}^\infty([0, 1] \times \Omega)$ ,

$$\iint_{[0,1] \times \Omega} \rho(\partial_t \Phi + v \cdot \nabla \Phi) dt dx = \int_{\Omega} [\rho_1 \Phi(1) - \rho_0 \Phi(0)] dx \quad (2.1.2)$$

Voir un exemple de solution de l’équation de continuité sur la figure 2.2.

### 2.1.3 Transport optimal et la distance de Wasserstein

#### Définition 2.1.3 (*Transport*)

Un transport entre  $\rho_0$  et  $\rho_1$  est un couple  $(\rho(t, x), v(t, x))$  solution de l’équation de continuité (2.1.2). L’ensemble des transports  $(\rho, v)$  sera noté

$$C(\rho_0, \rho_1).$$

#### Remarque 2.1.4

Pour  $\rho_0$  et  $\rho_1$  deux densités de probabilité, l’ensemble  $C(\rho_0, \rho_1)$  n’est pas vide. On trouvera



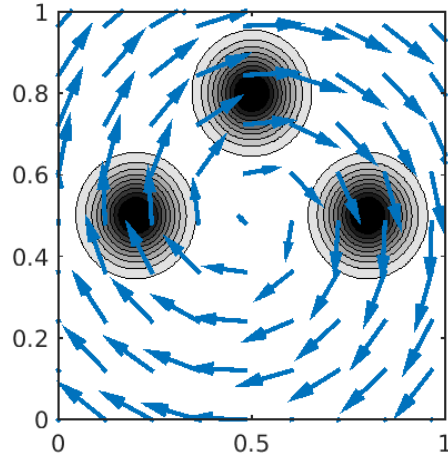


FIGURE 2.2 – Exemple de transport. Les flèches représentent le champ de vitesse  $v$ . La condition initiale  $\rho_0$  est à gauche : c'est une gaussienne centrée en  $(0.2, 0.5)$ . Sous l'action de  $v$ , via l'équation de transport, la gaussienne se déplace pour passer à  $(0.5, 0.8)$  à  $t = 1/2$  et arriver sur  $\rho_1$  à droite à  $t = 1$ .

toujours un champ de vitesse  $v$  qui transporte une densité sur l'autre.

Par exemple, si  $M > \rho_0(x) > \epsilon$  et  $M > \rho_1(x) > \epsilon$  avec  $M, \epsilon \in \mathbb{R}_+^*$ , alors on peut prendre le chemin  $\rho(t) = \rho_0 + t(\rho_1 - \rho_0)$ . Le champ de vitesse associé  $v(t, x)$  peut être pris égal à  $\nabla\Phi$ , avec  $\Phi(t, x)$  résolvant pour tout temps le problème elliptique :

$$\begin{aligned} -\operatorname{div}(\rho(t)\nabla\Phi(t)) &= \rho_1 - \rho_0 && \text{dans } \Omega \\ \frac{\partial\Phi}{\partial\vec{n}} &= 0 && \text{sur } \partial\Omega \end{aligned}$$

qui a bien une solution car l'intégrale de  $\rho_1 - \rho_0$  sont les mêmes [Allaire, 2007, Sec. 5.2.3.]. Lorsque  $\rho_0$  et  $\rho_1$  ne sont pas bornés par  $M$  ou s'annulent, il n'y a pas nécessairement de solution au problème elliptique, mais on affirme quand même qu'il y a toujours existence d'un chemin  $(\rho, v)$ . Il y a notamment le chemin optimal, qu'on détaillera après et qui existe d'après le théorème de Brenier (cf. le théorème 2.3.1).

Pour chaque couple  $(\rho, v)$  de  $C(\rho_0, \rho_1)$ , on peut associer un coût de déplacement. Classiquement on utilise l'énergie cinétique du transport, définie par

$$E(\rho, v) = \frac{1}{2} \iint_{[0,1] \times \Omega} \rho |v|^2 dt dx.$$

Transporter de manière optimale une densité  $\rho_0$  sur une autre  $\rho_1$  consiste à trouver le transport minimisant le coût de déplacement. C'est le problème du transport optimal. La valeur minimale définit ainsi le carré de la distance de Wasserstein qui vaut :

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \inf_{(\rho, v) \in C(\rho_0, \rho_1)} E(\rho, v). \quad (2.1.3)$$

Cette formulation a été introduite dans [Benamou et Brenier, 2000] et s'appelle formulation Benamou-Brenier du transport optimal. Nous introduirons dans la section 2.3 d'autres définitions équivalentes

du transport optimal et de la distance de Wasserstein.

**Proposition 2.1.5**

$\mathcal{W}_2$  est une distance sur  $\mathcal{P}(\Omega)$ .

C'est un résultat que l'on peut retrouver dans [Villani, 2003, Th. 7.3.]. Ceci va nous être utile pour doter  $\mathcal{P}(\Omega)$  d'une métrique, et définir par exemple une distance entre des images. D'autres distances existent sur  $\mathcal{P}(\Omega)$ , par exemple simplement la distance euclidienne  $\mathcal{L}^2$  dont le carré est défini par :

$$\mathcal{L}^2(\rho_0, \rho_1)^2 = \|\rho_0 - \rho_1\|_2^2 = \int_{\Omega} (\rho_0(x) - \rho_1(x))^2 dx$$

qui correspond à prendre un coût de déplacement égal à :

$$E(\rho, v) = \iint_{[0,1] \times \Omega} |\partial_t \rho|^2 dt dx.$$

Le choix de ce coût pour la distance euclidienne peut se démontrer en utilisant une inégalité de Jensen dans l'intégrale en temps.

Les caractéristiques spécifiques à la distance  $\mathcal{W}_2$  sont explicitées dans la section suivante à travers quelques exemples. On fera notamment une distinction entre cette distance et la distance euclidienne, et dans la suite du manuscrit nous essaierons d'exploiter les avantages de la distance de Wasserstein pour l'assimilation de données.

**Remarque 2.1.6**

Il existe beaucoup d'autres coûts de transport ; par exemple la distance  $\mathcal{W}_p$  ( $p \geq 2$ ) définie par :

$$\mathcal{W}_p(\rho_0, \rho_1)^p := \inf_{(\rho, v) \in C(\rho_0, \rho_1)} \frac{1}{p} \iint_{[0,1] \times \Omega} \rho |v|^p dt dx.$$

Ces distances sont équivalentes pour  $p > 1$  et nous utiliserons la distance  $\mathcal{W}_2$  par choix. On peut aussi ajouter d'autres termes pour pénaliser les déformations, formant un transport optimal plus réaliste, mais ces transports-ci ne constituent pas toujours une distance. Voir quelques exemples dans [Hug et al., 2015].

## 2.2 Quelques caractéristiques de la distance de Wasserstein

Depuis que des résultats théoriques et des algorithmes numériques performants ont été écrits (voir la partie 2.6 et l'annexe A pour ces algorithmes), le nombre d'applications du transport optimal ne fait que croître. Ainsi, nous allons exposer l'intérêt de la distance de Wasserstein en pointant quelques-unes de ses caractéristiques et montrer leur importance dans les applications à la recherche.

### 2.2.1 Une distance sur la position de l'information

Tout d'abord, la distance de Wasserstein est une métrique sur la position de l'information. Pour illustrer ce point-ci, on expose deux résultats significatifs. Le premier révèle que la distance est sensible aux déplacements, le deuxième qu'elle n'est pas sensible au bruit haute fréquence et on le verra, à la forme exacte décrite par l'information.

**Proposition 2.2.1 (Sensibilité au déplacement)**

Soit  $\rho_0 \in \mathcal{P}(\Omega)$  à support non égal à  $\Omega$ , et  $\rho_1: x \mapsto \rho_0(x - \mathbf{d})$  avec  $\mathbf{d}$  un vecteur de déplacement indépendant de  $x$ , et de telle sorte que  $\rho_1 \in \mathcal{P}(\Omega)$ . Alors

$$\mathcal{W}_2(\rho_0, \rho_1) = \frac{1}{\sqrt{2}}|\mathbf{d}|$$

Pour n'importe quel  $\mathbf{d}$  tel que  $\rho_0$  et  $\rho_1$  sont à support disjoints,  $\mathcal{L}^2(\rho_0, \rho_1) = \sqrt{2}\|\rho_0\|_2$ , qui est indépendant de  $\mathbf{d}$ . Ainsi, lorsque les supports sont disjoints, la distance euclidienne n'est pas sensible au déplacement. On verra même dans un exemple que les supports s'intersectent et que la distance euclidienne ne se comporte pas de manière régulière.

**Démonstration :** soit  $\rho(t) = \rho_0(x - t\mathbf{d})$  et  $v = \mathbf{d}$ . Le couple  $(\rho, v)$  appartient à  $C(\rho_0, \rho_1)$  et donc constitue bien un chemin admissible. On pourra montrer plus tard à l'aide du Théorème de Brenier 2.3.1 que c'est le couple optimal. Ainsi, la distance de Wasserstein s'écrit simplement

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \frac{1}{2} \iint_{[0,1] \times \Omega} \rho |v|^2 dt dx = \frac{1}{2} \iint_{[0,1] \times \Omega} \rho_0(x - t\mathbf{d}) |\mathbf{d}|^2 dt dx = \frac{|\mathbf{d}|^2}{2} \quad \blacksquare$$

Une autre caractéristique est la non-sensibilité au bruit haute fréquence. En effet, on a la proposition suivante tirée de [Villani, 2003, Th. 7.12].

**Théorème 2.2.2**

Soit  $\rho^n$  une suite de  $\mathcal{P}(\Omega)$ ,  $\mathcal{W}_2(\rho^n, \rho) \rightarrow 0$ , alors si et seulement si  $\rho^n$  converge faiblement vers  $\rho$  et que les moments d'ordre 2 de  $\rho^n$  convergent vers ceux de  $\rho$ .

La convergence faible implique (notamment) la convergence à un bruit haute fréquence près, comme on peut le voir dans l'exemple suivant, où l'on ajoute un bruit qui sera quasi-invisible pour la distance de Wasserstein.

**Exemple 2.2.3 (Non-sensibilité au bruit haute fréquence)**

Soit  $\Omega = [0, 1]$ ,  $\rho_0 \in \mathcal{P}(\Omega)$  tel que  $\rho_0 > m > 0$ , et

$$\rho_1^n(x) = \rho_0(x) + \alpha \cos(2n\pi x)$$

avec  $\alpha < m$ , de telle sorte que  $\rho_1^n \in \mathcal{P}(\Omega)$ ;  $n \geq 1$  est un entier représentant la fréquence du bruit. Alors

$$\mathcal{W}_2(\rho_0, \rho_1^n) \leq \frac{C}{n}$$

avec  $C$  une constante dépendant de  $\alpha$  et de  $\rho_0$  mais pas de  $n$ . Ainsi, quand  $n \rightarrow \infty$ , la distance de Wasserstein tend vers 0.

Ainsi, pour un  $n$  grand, manifestant un bruit de haute fréquence, la distance de Wasserstein devient petite. Par comparaison,  $\mathcal{L}^2(\rho_0, \rho_1^n) = \frac{\alpha}{\sqrt{2}}$  quel que soit  $n \geq 1$ .

**Démonstration :** Soit  $\rho(t, x) = \rho_0(x) + \alpha t \cos(2n\pi x)$ . Si l'on pose

$$v = -\frac{\alpha}{2\pi n \rho} \sin(2\pi n x),$$

alors le couple  $(\rho, v)$  fait partie des chemins admissibles  $C(\rho_0, \rho_1^n)$ , puisque notamment  $\partial_t \rho + \operatorname{div}(\rho v) = 0$ . Donc

$$\mathcal{W}_2(\rho_0, \rho_1^n)^2 \leq \frac{1}{2} \iint_{[0,1] \times [0,1]} \rho |v|^2 dt dx$$

ce qui donne après remplacement de la valeur de  $v$

$$\mathcal{W}_2(\rho_0, \rho_1^n)^2 \leq \frac{\alpha^2}{8\pi^2 n^2} \iint_{[0,1] \times [0,1]} \frac{\sin(2\pi n x)^2}{\rho} dt dx$$

et par majoration de l'intégrale, comme le sinus est plus petit que 1 et que  $\rho(t) \geq m - \alpha > 0$ ,

$$\mathcal{W}_2(\rho_0, \rho_1^n)^2 \leq \frac{\alpha^2}{4\pi^2 n^2 (m - \alpha)}$$

ce qui démontre le résultat. ■

Ces deux résultats illustrent que la distance de Wasserstein est sensible à la localisation de l'information mais pas aux détails de l'information donc à sa forme exacte prise. Cette propriété présente quelques applications dans beaucoup de champs de la science. Tout d'abord, on peut citer son importance dans les mathématiques théoriques et dans l'étude des équations différentielles. Ensuite, il y a les applications en imagerie et pour les problèmes inverses. Enfin, nous verrons une autre caractéristique conséquente aux propositions que l'on vient d'énoncer et qui nous sera très utile.

**Applications aux mathématiques théoriques.** L'analyse utilise beaucoup le fait que la distance de Wasserstein traduit des propriétés physiques, notamment du fait que la distance de Wasserstein métrise la convergence faible, voir le théorème 2.2.2 ci-dessus. Le bruit est très présent en physique. C'est pourquoi la distance de Wasserstein et le transport optimal sont utilisés par exemple pour l'équation de Boltzmann qui décrit l'évolution d'un gaz [Tanaka, 1978], pour les équations semi-géostrophiques, voir [Benamou et Brenier, 1998, Cullen et Gangbo, 2001] ou en astrophysique pour la reconstruction de l'état primitif de l'Univers [Brenier et al., 2003].

Récrire les équations à l'aide de gradients de Wasserstein permet de les réinterpréter physiquement. Ceci s'appelle les flots gradients. Par exemple, pour l'équation de la chaleur,

$$\partial_t \rho + \Delta \rho = 0$$

qui modélise la propagation de la chaleur dans un milieu, cette équation peut se récrire

$$\rho'(t) = -\nabla E(\rho(t))$$

avec  $\nabla$  le gradient de Wasserstein (que nous définirons plus loin) et  $E(\rho) = -\int_{\Omega} \rho \log(\rho)$  est l'entropie associée à  $\rho$ . Cela signifie que les solutions de l'équation de la chaleur sont les courbes qui font le plus décroître l'entropie pour la distance de Wasserstein. L'exemple précurseur des flots gradients de Wasserstein est de [Jordan et al., 1998].

De plus, le transport optimal permet de prouver de manière courte et élégante quelques inégalités fonctionnelles [Cordero-Erausquin et al., 2004, Agueh, 2008, Otto et Villani, 2000, Villani, 2003, Villani, 2008].

**Retrouver des images dans une base de données.** Le fait que le transport optimal mesure le lieu de l'information sert aussi pour les bases de données d'images, étant donné que dans une image, la position des structures compte. Par exemple, on voudrait retrouver les images d'une base de données les plus proches d'une image cible. La notion de proximité est à établir. Dans [Rubner et al., 2000] les auteurs utilisent la distance  $\mathcal{W}_1$  qu'ils utilisent sur les histogrammes de couleur des images. Les résultats sont visuellement meilleurs, cf. la figure 2.3, où les résultats obtenus sont comparés avec une méthode classique.

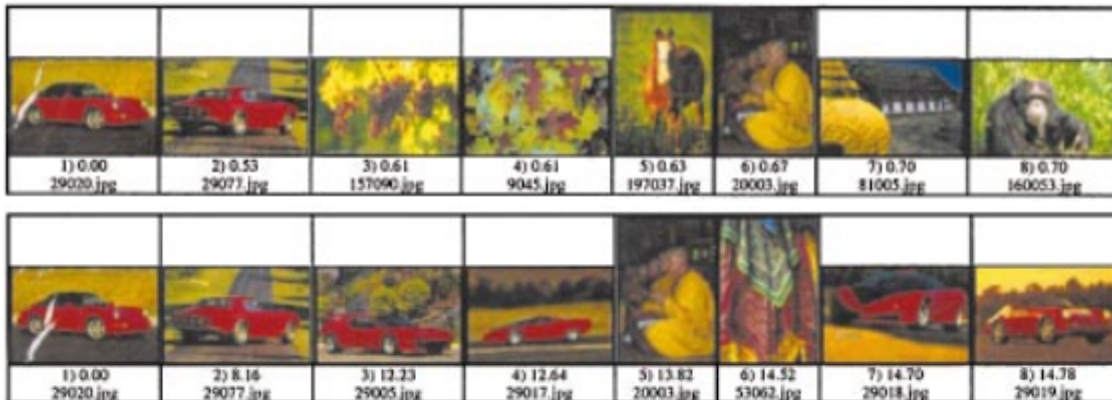


FIGURE 2.3 – Les huit éléments d’une base de données les plus proches de la voiture rouge de gauche, (en haut) pour une distance  $L^1$  sur les histogrammes de couleur, (en bas) pour la distance  $\mathcal{W}_1$  qu’ils appellent aussi EMD (Earthe Mover’s distance). Tiré de [Rubner et al., 2000].

**Application aux problèmes inverses.** On a vu dans la proposition 2.2.1 que la distance de Wasserstein est sensible aux déplacements, contrairement à la distance  $\mathcal{L}^2$ . Dans la figure 2.4, on compare l’évolution des valeurs des distances euclidienne et de Wasserstein entre deux fonctions  $\rho_0$  et sa translatée  $\rho_1 = \rho_0(\cdot - d)$ , en fonction de  $d$ . On remarque que pour une fonction oscillante notamment, la valeur de la distance de Wasserstein en fonction de  $d$  se comporte beaucoup mieux que pour celle de la distance euclidienne, il y a notamment convexité par rapport au déplacement. C’est cette propriété de convexité que les auteurs de [Engquist et Froese, 2013] proposent d’exploiter pour comparer des ondes sismiques notamment, qui sont des fonctions oscillantes et qui se déplacent dans le temps (avance/retard) en fonction de la conductivité du matériau. Pour retrouver cette conductivité, on pourrait écrire une fonction coût, dépendant de la conductivité, à l’aide de la distance de Wasserstein, qui serait donc convexe, et facilement minimisable. On retraitera cet exemple plus en détail dans le chapitre 3.

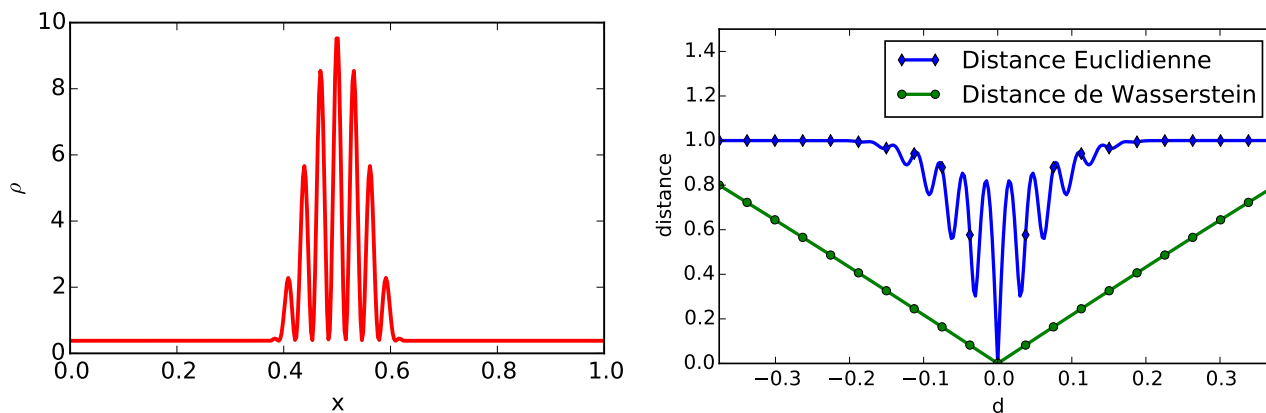


FIGURE 2.4 – Avec le  $\rho_0$  tracé à gauche et  $\rho_1 = \rho_0(\cdot - d)$ , on trace à droite en fonction de  $d$  les valeurs de  $\mathcal{L}^2(\rho_0, \rho_1)$  et de  $\mathcal{W}_2(\rho_0, \rho_1)$  (normalisées). On remarque notamment que la distance est convexe par rapport à  $d$  pour la seconde distance alors qu’elle ne l’est pas pour la première.

**Autres applications.** Comme autre application on peut trouver le tracé d'un contour d'un objet en minimisant une fonctionnelle basée sur la distance de Wasserstein [Peyré et al., 2012] ou le transfert de couleur donnant des résultats visuellement intéressants voir [Piti'e et al., 2005] ou plus récemment [Ferradans et al., 2014, Rabin et al., 2011], cf. la figure 2.5.

Le transport optimal trouve également beaucoup d'applications en imagerie médicale comme aide au personnel soignant. Il peut servir à détecter les zones d'activation neuronales en comparant différents cerveaux [Gramfort et al., 2015] ou la détection de cerveaux sclérosés [Saumier et al., 2010], trouver le champ de déplacement de l'hippocampe [Bouharguane et al., 2014] en calculant le transport entre deux images de cerveau prises à des moments différents, suivre les tumeurs grossissantes [Lombardi et Maitre, 2015], simuler le transport de traceurs dans le système glymphatique [Ratner et al., 2015].



(a) Image d'origine

(b) Cible

(c) Résultat du transfert

FIGURE 2.5 – Exemple de transfert de couleur, où l'on tente de remplacer les couleurs de l'image (a) par celles de l'image (b). Le résultat est l'image (c). Tiré de [Rabin et al., 2011]

### 2.2.2 Interpoler à l'aide de la distance de Wasserstein

#### Définition 2.2.4 (*Barycentre*)

Un barycentre de deux ou plusieurs éléments  $x_i$  pour une distance  $d$  donnée est la solution du problème de minimisation

$$\min_x \sum_i \omega_i d(x, x_i)^2$$

avec  $\omega_i \geq 0$  des poids donnés.

Le barycentre (avec des poids égaux) pour la distance de Wasserstein, de deux densités localisées  $\rho_0$  et  $\rho_1$  est une troisième densité positionnée entre les deux, voir la figure 2.6, où l'on montre aussi le barycentre pour la distance euclidienne, défini comme étant la moyenne des deux,

$$\frac{\rho_0 + \rho_1}{2}.$$

Ainsi, le barycentre pour la distance de Wasserstein a tendance à conserver la forme de ce qu'il interpole. On montre sur la figure 2.7 un application entre les barycentres pour chacune des deux distances, pour retrouver des formes soumises à des déformations.

Cette propriété d'interpolation est intéressante puisque le barycentre change la position de ce qu'il interpole. C'est une propriété que nous utiliserons plus tard ; mais avant on montre son utilité dans quelques applications en imagerie.

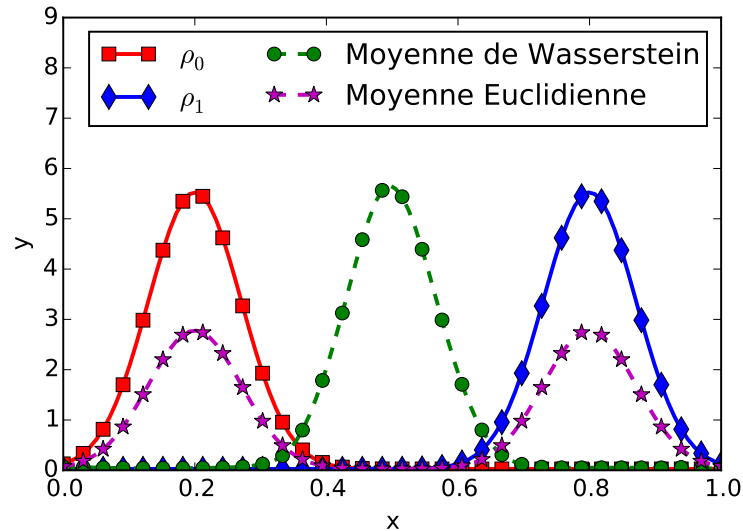


FIGURE 2.6 – Moyenne des deux gaussiennes  $\rho_0$  et  $\rho_1$  pour la distance de Wasserstein et la distance Euclidienne.

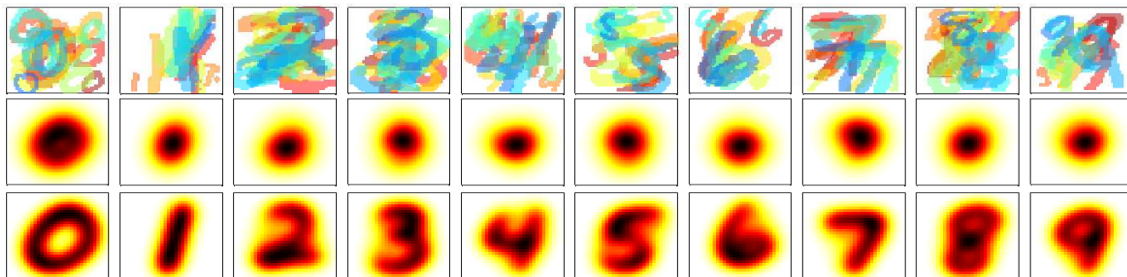


FIGURE 2.7 – Pour les 10 chiffres de taille et de formes différentes (première ligne), on calcule les barycentres  $\mathcal{L}_2$  après recentrage (deuxième ligne) et  $\mathcal{W}_2$  (troisième ligne). Source : [Doucet et Cuturi, 2014]

**L’interpolation d’images et barycentres.** Cette propriété a quelques applications, comme la création de nouvelles textures visuellement intéressantes par interpolation, voir [Rabin et al., 2011] où une distance de Wasserstein légèrement simplifiée a été utilisée, et cf. figure 2.8. On peut aussi utiliser la distance de Wasserstein pour reconstruire des formes manquantes : en projetant des densités “trouées” sur l’espace des interpolants d’une base de données, on arrive à reconstruire l’information manquante [Bonneel et al., 2016], voir figure 2.9. On peut aussi citer le rééquilibrage d’histogrammes [Delon, 2004, Guillemot et Delon, 2016] basé sur l’interpolation au sens de Wasserstein des histogrammes de niveau de gris. Cela recrée une interpolation visuellement meilleure. Une application est la réparation de films, cf. [Delon et Desolneux, 2010], pour réduire les effets de scintillements (locaux) que l’on peut trouver dans les vieux films par exemple. Ceci est réalisé en moyennant l’histogramme des niveaux de gris de chaque image du film avec ceux des images voisines.

**Remarque 2.2.5 (Interpolation pour les fonctions à support égal à  $\Omega$ )**

Considérons  $\Omega = [0, 1]$  et deux fonctions  $\rho_0$  et  $\rho_1$  tels que  $\rho_0 \approx 1$  et  $\rho_1 \approx 1$ . Alors le barycentre des deux pour la distance de Wasserstein est proche du barycentre pour la distance euclidienne.



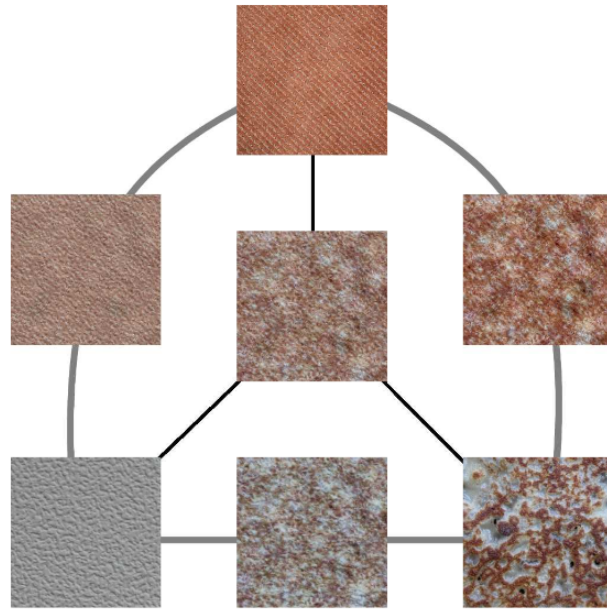
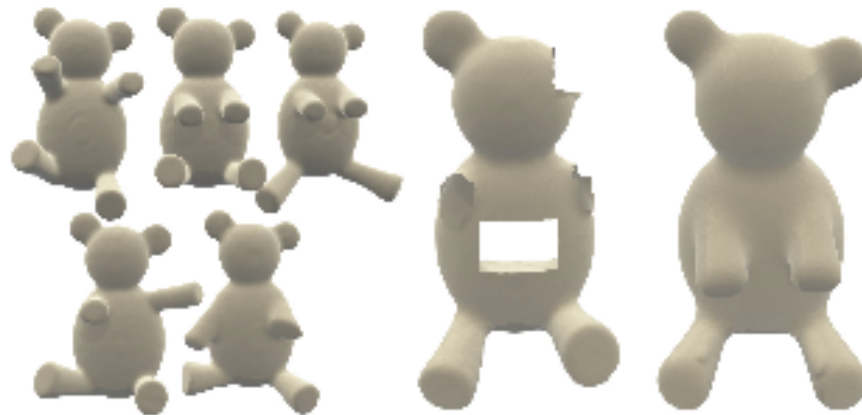


FIGURE 2.8 – Les trois textures originales sont dans les sommets du triangle, en haut au milieu et en bas à droite et à gauche. Les interpolations de texture sont sur les côtés et au centre du triangle. Tiré de [Rabin et Peyré, 2011].



(a) Base de données    (b) Élément à reconstruire    (c) Résultat

FIGURE 2.9 – On dispose d’une base de données (a) et d’un ourson dont il manque des parties (b). En projetant cet ourson sur l’espace des interpolées (pour la distance de Wasserstein) de (a), on retrouve l’élément (c). Image tirée de [Bonneel et al., 2016].

*Nous montrons ce point-ci dans la figure 2.10. En général, pour que les structures du barycentre se situent entre les structures des densités interpolées, il faut que ces structures soient des masses “localisées”, c’est-à-dire entourées d’une zone où la densité est quasi-nulle.*



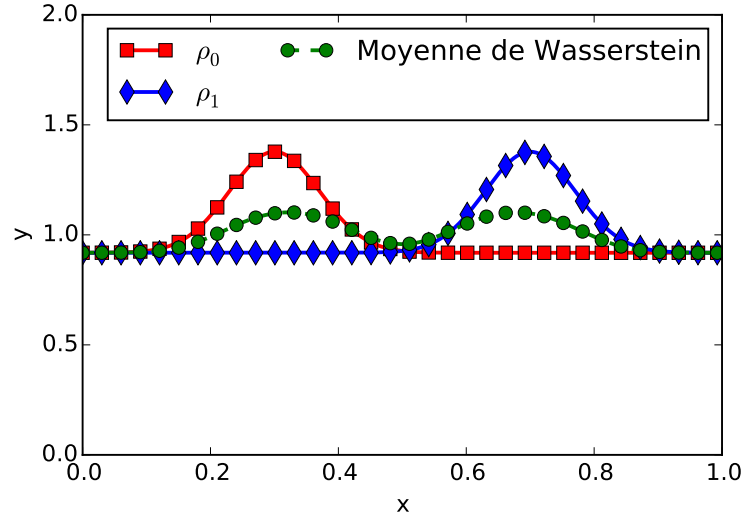


FIGURE 2.10 – Moyenne de Wasserstein des deux densités  $\rho_0$  et  $\rho_1$  proches de 1 (voir l’axe des ordonnées).

## 2.3 Autres définitions de la distance de Wasserstein

Dans cette partie on prendra  $\rho_0$  et  $\rho_1$  deux éléments de  $\mathcal{P}(\Omega)$ . On rappelle que la formulation Benamou-Brenier définit le carré de la distance de Wasserstein,

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \inf_{(\rho, v) \in C(\rho_0, \rho_1)} E(\rho, v) = \frac{1}{2} \iint_{[0,1] \times \Omega} \rho |v|^2 dt dx \quad (\text{BB})$$

avec

$$C(\rho_0, \rho_1) := \left\{ (\rho, v) \text{ tel que } \begin{array}{l} \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \rho(t=0) = \rho_0, \rho(t=1) = \rho_1, \\ \rho v \cdot \vec{n} = 0 \text{ p.p. sur } [0, 1] \times \partial\Omega \end{array} \right\}$$

où les équations sont à prendre au sens faible. Il existe d’autres écritures de la distance de Wasserstein qui sont les plus couramment introduites dans les ouvrages de référence du transport optimal puisque historiquement plus anciennes. Nous allons les présenter dans cette partie. L’équivalence entre les différentes écritures de la distance ajoute de la richesse à la compréhension du transport optimal. En particulier, nous les utiliserons toutes dans la suite de ce manuscrit pour prouver différents résultats ou pour la construction de schémas numériques.

Nous écrirons donc les équivalences entre les différentes formes du transport optimal. Ces équivalences sont bien connues. Cependant, les liens entre les différents éléments optimaux sont décrits de manière succincte et éparse dans la littérature, c’est pourquoi nous avons souhaité les regrouper proprement dans ce manuscrit. On notera notamment l’élément central de ces éléments optimaux qui est le potentiel de Kantorovich.

### 2.3.1 Le problème de Monge

En 1781 Gaspard Monge introduit le problème du transport optimal dans un traité appelé *Mémoires sur les déblais et les remblais* [Monge, 1781]. G. Monge se pose le problème suivant : comment remplir un trou avec un tas de sable, en minimisant l’énergie totale pour le déplacer ?

On représente le tas de sable par une densité  $\rho_0(x)$ , le trou par une densité  $\rho_1(y)$ <sup>1</sup>, et le déplacement par une fonction  $T(x)$  qui déplace la quantité de sable présente en  $x$  vers le point  $y = T(x)$ , voir figure 2.11. Le coût du déplacement associé à  $T$  est :

$$\int_{\Omega} c(x, T(x)) \rho_0(x) dx.$$

Le terme  $c(x, y)$  est le coût de déplacement d'une unité de sable de  $x$  vers  $y$ . Il est courant de prendre le coût quadratique  $c(x, y) = |x - y|^2/2$ . Ainsi, on cherche le déplacement (ou le plan de transport)  $T$  optimisant ce coût.

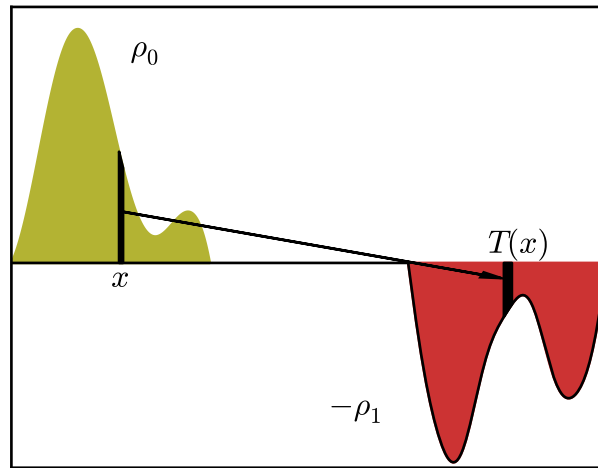


FIGURE 2.11 – Problème de Monge entre le tas de sable à gauche représenté par  $\rho_0(x)$  et le trou à droite représenté par  $-\rho_1(y)$ . Le plan de transport  $T(x)$  indique la destination du sable contenu en  $x$ . La valeur de  $\rho_0$  en  $x$  ne doit pas forcément être égale à la valeur de  $\rho_1$  en  $T(x)$  ; grâce aux effets de compression/dilatation locale de  $T$  on permet au sable de s'étaler ou de se concentrer en  $T(x)$ .

Il y a plusieurs contraintes sur les données :

- Le tas de sable et le trou doivent avoir le même volume, qu'on fixe égal à 1. De plus les fonctions qui les représentent doivent être positives. Ainsi on retrouve que la contrainte sur  $\rho_0$  et  $\rho_1$  est :

$$\rho_0, \rho_1 \in \mathcal{P}(\Omega).$$

- Le déplacement  $T$  est aussi assujéti à une contrainte, la contrainte de transport. En effet, il doit déplacer le tas dans le trou, ce qu'on note  $T\#\rho_0 = \rho_1$ . On va écrire ceci mathématiquement. On ne demande pas que la masse envoyée sur  $T(x)$  soit exactement la même que celle provenant de  $x$ . Ainsi on peut avoir  $\rho_0(x) \neq \rho_1(T(x))$ . Plutôt, on écrit les choses ainsi : soit  $A \subset \Omega$  un ensemble quelconque. Tout le sable qui a été transporté sur  $A$  est exactement tout le sable qui était en  $T^{-1}(A)$  :

$$\forall A \subset \Omega, \int_{T^{-1}(A)} \rho_0(x) dx = \int_A \rho_1(y) dy. \quad (2.3.1)$$

1. Le réel problème que se posait Monge était celui-ci mais dans le cas où  $\rho_0$  et  $\rho_1$  valent 1 sur un ensemble et 0 ailleurs.

Pour des plans de transport  $T$  différentiables, on peut effectuer un changement de variable dans (2.3.1) et la condition de transport est équivalente à :

$$\rho_1(T(x)) |\det(\nabla T)|(x) = \rho_0(x). \quad (2.3.2)$$

Pour résumer, le problème de Monge est :

$$\begin{aligned} & \underset{T}{\text{minimiser}} && \frac{1}{2} \int_{\Omega} |x - T(x)|^2 \rho_0(x) \, dx \\ & \text{sous la contrainte :} && T\#\rho_0 = \rho_1 \text{ (voir (2.3.1) ou (2.3.2))} \end{aligned} \quad (\text{M})$$

Il existe un résultat d'existence d'un plan de transport  $T$  minimiseur. C'est le théorème de Brenier [Brenier, 1991] :

**Théorème 2.3.1 (Théorème de Brenier)**

Soient  $\rho_0, \rho_1 \in \mathcal{P}(\Omega)$ . Il existe une fonction scalaire  $\Psi(x)$ , convexe, telle que  $T = \nabla\Psi$  est minimiseur de (M).

**Théorème 2.3.2 (Théorème de Benamou-Brenier [Benamou et Brenier, 2000])**

Les deux problèmes (BB) et (M) sont équivalents, c'est-à-dire que les valeurs des minima sont égales (au carré de la distance de Wasserstein). De plus, le  $(\rho, v) \in C(\rho_0, \rho_1)$  optimum pour (BB) et le  $T$  optimal pour (M) sont liés par le fait que  $v(t, x)$  résout l'équation de Burgers

$$\partial_t v + v \cdot \nabla v = 0$$

de condition initiale

$$v(0, x) = T(x) - x.$$

L'ébauche de la démonstration manipule l'équation de continuité pour passer d'une formulation à l'autre. Elle n'est pas essentielle pour la compréhension du reste du manuscrit. C'est une ébauche puisqu'elle suppose un peu de régularité sur les densités. Elle est tirée de [Benamou et Brenier, 2000].

**Démonstration :** La preuve s'appuie sur la résolution de l'équation de continuité par les caractéristiques. Soit  $v$  une vitesse, si  $\rho(t, x)$  est solution de l'équation de continuité

$$\partial_t \rho + \operatorname{div}(\rho v) = 0,$$

alors il existe un plan de transport  $T_t(x)$  (la caractéristique) tel que  $\rho$  est donné explicitement par

$$\rho(t, T_t(x)) |\det(\nabla T_t(x))| = \rho_0(x) \quad (2.3.3)$$

qu'on note aussi  $T_t\#\rho_0 = \rho(t)$ , avec  $T_t(x)$  la caractéristique de pied  $x$ , définie par l'équation

$$\frac{\partial T_t(x)}{\partial t} = v(t, T_t(x)) \text{ avec } T_0(x) = x. \quad (2.3.4)$$

Dans le problème (BB), l'énergie cinétique se réécrit avec ces caractéristiques à l'aide d'un changement de variable  $x = T_t(y)$ , ce qui fait disparaître la vitesse. En effet, soit  $(\rho, v) \in C(\rho_0, \rho_1)$  et soit la

caractéristique  $T_t$  liée à  $v$ , alors l'énergie cinétique se réécrit

$$\begin{aligned}
E(\rho, v) &= \frac{1}{2} \iint_{[0,1] \times \Omega} \rho |v|^2 dt dx \stackrel{x=T_t(y)}{=} \frac{1}{2} \iint_{[0,1] \times \Omega} \rho(t, T_t) |v(t, T_t)|^2 |\det(\nabla T_t)| dt dy \\
&\stackrel{\substack{(2.3.3) \\ (2.3.4)}}{=} \frac{1}{2} \iint_{[0,1] \times \Omega} \rho_0(y) |\partial_t T_t(y)|^2 dt dy \\
&\stackrel{\text{Jensen}}{\geq} \frac{1}{2} \int_{\Omega} \rho_0(y) \left| \int_{[0,1]} \partial_t T_t(y) dt \right|^2 dy \\
&= \frac{1}{2} \int_{\Omega} \rho_0(y) |T_1(y) - y|^2 dy.
\end{aligned} \tag{2.3.5}$$

Et comme  $T_1$  vérifie la condition de transport  $T_1 \# \rho_0 = \rho_1$ , on a prouvé en prenant l'infimum sur tous les plans de transport  $T$  que

$$\forall (\rho, v) \in C(\rho_0, \rho_1), \quad E(\rho, v) \geq C_M \tag{2.3.6}$$

en notant  $C_M$  la valeur du problème de Monge (M). On rappelle que le problème (BB) consiste à prendre l'infimum (noté  $C_{BB}$ ) de l'énergie cinétique sur les  $(\rho, v)$  appartenant à  $C(\rho_0, \rho_1)$ . Ainsi, en prenant l'infimum dans (2.3.6), on peut tomber sur deux cas : soit  $C_{BB} > C_M$ , soit  $C_{BB} = C_M$ . On va montrer qu'on est dans ce deuxième cas en exhibant un couple  $(\rho, v)$  dont l'énergie cinétique égale la valeur du problème de Monge.

En reprenant le calcul (2.3.5), on a en fait en reprenant les notations que l'égalité

$$E(\rho, v) = \frac{1}{2} \int_{\Omega} \rho_0(y) |T_1(y) - y|^2 dy.$$

n'est possible que dans le cas d'égalité dans l'inégalité de Jensen, ce qui arrive lorsque  $\partial_t T_t(x)$  est une constante du temps ; donc  $T_t(x) = x + t(T_1(x) - x)$ . Ceci se produit lorsque la vitesse  $v$  associée résout l'équation de Burgers

$$\partial_t v + v \cdot \nabla v = 0$$

de condition initiale  $v(0, x) = T_1(x) - x$ .

Ainsi, soit  $T$  le transport optimal entre  $\rho_0$  et  $\rho_1$ , celui qui résout le problème de Monge (M). Soit  $v(t)$  un champ de vitesse qui résout l'équation de Burgers de condition initiale  $v(0, x) = T(x) - x$ . Ce champ  $v(t)$  existe pour  $t \in [0, 1]$  car la caractéristique associée  $T_t(x) = (1-t)x + tT(x)$  est toujours inversible (voir le théorème de Brenier 2.3.1 qui dit que  $T$  a les propriétés suffisantes pour : c'est le gradient d'une fonction convexe). Soit  $\rho(t)$  la densité résolvant l'équation de continuité avec cette vitesse  $v$ , et de condition initiale  $\rho_0$ . Par construction, ce couple est d'énergie cinétique égale à la valeur du problème de Monge. Il reste à montrer que ce couple  $(\rho, v)$  appartient à  $C(\rho_0, \rho_1)$  :

- premièrement,  $\rho(t)$  satisfait l'équation de continuité et  $\rho(0) = \rho_0$  ;
- ensuite,  $\rho(1) = \rho_1$ . En effet, comme  $\rho$  vérifie l'équation de continuité, (2.3.3) est vérifié avec  $T_t$  la caractéristique. Donc on a à  $t = 1$  :

$$\rho(1, T_1) |\det(\nabla T_1)| = \rho_0.$$

De plus,  $T$  est le transport optimal entre  $\rho_0$  et  $\rho_1$ , et vérifie notamment  $T \# \rho_0 = \rho_1$  :

$$\rho_1(T) |\det(\nabla T)| = \rho_0.$$

En utilisant que  $T_1 = T$  et en divisant une équation par l'autre, on trouve que  $\rho(1) = \rho_1$ .

- Enfin,  $\rho v \cdot \vec{n} = 0$  presque partout sur le bord du domaine et pour presque tout temps. En effet, ceci se base sur (i) le fait que la caractéristique s'écrit  $T_t(x) = tT(x) + (1-t)x$ , que (ii) le domaine  $\Omega$  est convexe, (iii)  $T(x) \in \Omega$  pour tout  $x$ . Par conséquent,  $T_t(x) \in \Omega$  pour tout temps. Ainsi, une particule ne sort jamais du domaine, ce qui s'exprime par le fait que  $v(t, x) \cdot \vec{n}(x) \leq 0$  sur  $\partial\Omega$ . Mais comme  $\rho_0$  et  $\rho_1$  sont de même masse, on a en fait que  $\rho(t, x) v(t, x) \cdot \vec{n}(x) = 0$  pour presque tout  $(t, x) \in [0, 1] \times \partial\Omega$ .

Donc ceci montre que  $C_{BB} = C_M$ . Le couple  $(\rho, v)$  ainsi construit est aussi le couple optimal pour le problème de Benamou-Brenier. ■

### Remarque 2.3.3

La différence entre le problème de Monge (M) et la formulation Benamou-Brenier (BB) est que l'on cherche un plan de transport  $T(x)$  à la place d'un chemin  $\rho(t, x)$  et d'un champ de vitesse  $v(t, x)$ . La dimension temporelle a disparu. Ce problème plus simple à écrire souffre d'un gros défaut : la contrainte de transport (2.3.2) est très difficile à gérer, notamment à cause du déterminant du gradient de  $T$ .

La démonstration du précédent théorème fait ressortir un point important du transport optimal : la caractérisation du chemin optimal  $(\rho, v)$ . La proposition suivante 2.3.4 permet premièrement d'écrire l'équation sur  $\rho$  plus simplement, et deuxièmement, de simplifier encore l'équation de Burgers sur  $v$ . Cette proposition est issue de [Villani, 2003, Ex.5.42].

### Proposition 2.3.4

Soit  $(\rho, v) \in C(\rho_0, \rho_1)$  le chemin optimal entre  $\rho_0$  et  $\rho_1$ , et  $T = \nabla\Psi$  le transport optimal. Alors :

1.  $\rho(t)$  s'écrit aussi :

$$\rho(t) = ((1-t)Id + t\nabla\Psi)\#\rho_0;$$

2. on a que  $v(t, x) = \nabla\phi(t, x)$ , avec  $\phi(t, x)$  solution de l'équation d'Hamilton-Jacobi

$$\partial_t\phi + \frac{|\nabla\phi|^2}{2} = 0, \quad \phi(0, x) = \Psi(x) - \frac{|x|^2}{2}. \quad (2.3.7)$$

Si bien que les équations du minimiseur du transport optimal  $(\rho, v = \nabla\phi)$  s'écrivent :

$$\begin{cases} \partial_t\rho + \operatorname{div}(\rho\nabla\phi) = 0 \\ \partial_t\phi + \frac{|\nabla\phi|^2}{2} = 0 \end{cases} \quad (2.3.8)$$

$$\rho(0) = \rho_0, \rho(1) = \rho_1, \quad \phi(0, x) = \Psi(x) - \frac{|x|^2}{2}.$$

**Démonstration :** Le premier point se démontre à l'aide des caractéristiques. Soit  $T_t$  la caractéristique associée à  $v$ . Comme  $v$  suit une équation de Burgers et que  $v(0, x) = T(x) - x$ , on a que

$$T_t(x) = x + tv(0, x) = (1-t)x + t\nabla\Psi(x)$$

Comme  $\rho$  suit une équation de continuité de vitesse  $v$ , on écrit sa solution comme dans (2.3.3) dans la preuve du théorème précédent,

$$\rho(t) = T_t\#\rho_0 = ((1-t)Id + t\nabla\Psi)\#\rho_0.$$

Pour le deuxième point, on prend  $\phi$  résolvant l'équation d'Hamilton-Jacobi (2.3.7). En appliquant l'opérateur  $\nabla$  à cette équation, on trouve que  $\nabla\phi$  résout également l'équation de Burgers avec la même condition initiale que  $v$ . Par unicité des solutions de l'équation de Burgers,  $\nabla\phi(t, x)$  et  $v(t, x)$  sont égaux. ■

### 2.3.2 Le problème de Kantorovich

Il existe encore d'autres versions de la distance de Wasserstein qui nous seront également utiles plus tard. La version (M) était celle proposée par Monge en 1781. Pour la résoudre Kantorovich en 1942 s'est aidé d'un problème plus simple, relâché, qui redonne aussi la distance de Wasserstein. Au lieu de chercher parmi les plans  $T$  qui transportent  $\rho_0$  sur  $\rho_1$ , le problème de Kantorovich consiste à trouver la mesure  $\pi(x, y)$  définie sur  $\Omega \times \Omega$ , de marginales  $\rho_0$  et  $\rho_1$ , qui minimise le déplacement quadratique.

Pour expliquer ce que cela signifie, on revient sur le problème du tas de sable et du trou à combler de Monge. Tout le sable contenu en  $x$  était transporté en  $T(x)$ . Le problème de Kantorovich consiste lui à trouver une mesure  $\pi(x, y)$  qui indique quelle quantité de sable est transportée de  $x$  vers  $y$ . Ainsi, on peut partager la masse contenue en  $x$ , ce qui n'était pas admis dans le problème de Monge puisque tout ce qui était en  $x$  allait en  $T(x)$ . Voir la Figure 2.12.

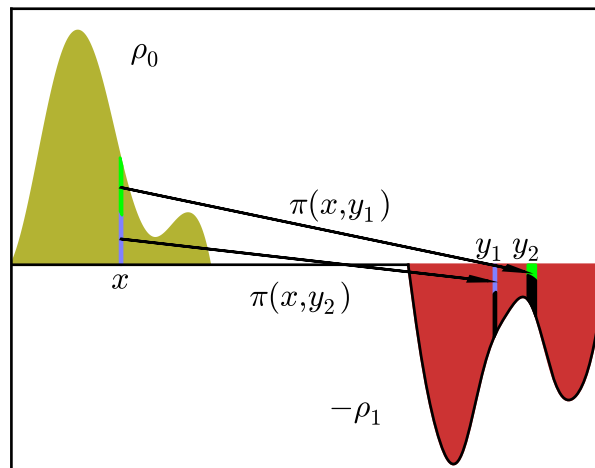


FIGURE 2.12 – On reprend l'exemple montré dans le problème de Monge. Ici,  $\pi$  partage la masse contenue en  $x$ , une partie va en  $y_1$  et l'autre en  $y_2$ .

La condition de transport de  $\rho_0$  sur  $\rho_1$  par  $\pi$  s'écrit ainsi :

- la totalité ce qui est pris de  $A \subset \Omega$  vaut  $\iint_{A \times \Omega} \pi(x, y) dx dy$ , mais aussi  $\int_A \rho_0(x) dx$  ;
- et la totalité ce qui est emmené sur  $B \subset \Omega$  vaut  $\iint_{\Omega \times B} \pi(x, y) dx dy$ , mais aussi  $\int_B \rho_1(y) dy$ .

Cela se récrit à l'aide de fonctions :

$$\forall \Phi, \Psi \in \mathcal{C}^\infty(\Omega), \iint_{\Omega \times \Omega} (\Phi(x) + \Psi(y)) d\pi(x, y) = \int_{\Omega} \Phi(x) \rho_0(x) dx + \int_{\Omega} \Psi(y) \rho_1(y) dy. \quad (2.3.9)$$

Ainsi on cherche  $\pi$  dans  $\Pi(\rho_0, \rho_1)$  avec

$$\Pi(\rho_0, \rho_1) = \left\{ \text{Les mesures } \pi \geq 0 \text{ vérifiant (2.3.9)} \right\}. \quad (2.3.10)$$

Ainsi, ce problème de Kantorovich consiste à résoudre

$$\inf_{\pi \in \Pi(\rho_0, \rho_1)} \frac{1}{2} \iint_{\Omega \times \Omega} |x - y|^2 d\pi(x, y). \quad (\text{K})$$

Le théorème suivant tiré de [Villani, 2003, Th. 2.12.ii] relie ce problème avec les autres problèmes de transport optimal décrits ci-dessus :

**Théorème 2.3.5**

Les problèmes de Monge (M) et de Kantorovich (K) sont équivalents. Le  $\pi$  optimal pour (K) est un Dirac le long d'une courbe, c'est-à-dire qu'il existe  $T$  tel que  $\pi(x, y) = \delta_{x=T(y)}$  ce qui nous permet de retrouver le problème de Monge. Ce  $T$  est le plan de transport optimal entre  $\rho_0$  et  $\rho_1$ , celui qui résout le problème de Monge (M).

On ne détaille pas la preuve qui ne nous apporte pas beaucoup ici, une ébauche est trouvée dans [Santambrogio, 2014] et une preuve complète dans [Villani, 2003, Th. 2.12]. La preuve repose néanmoins sur la dualité de Kantorovich introduite dans la section suivante.

Avec ce théorème, et par simple utilisation du Théorème précédent (2.3.2), le problème de Kantorovich redonne la distance de Wasserstein.

**2.3.3 Dualité de Kantorovich**

Enfin, parlons de la dualité de Kantorovich, introduit par L. Kantorovich en 1942 pour résoudre son problème, et qui sert à démontrer le Théorème 2.3.5. Le problème dual va aussi nous intéresser plus tard pour écrire la différentielle de la distance de Wasserstein. On peut montrer que ce problème est dual au sens de Fenchel-Rockafellar [Villani, 2003, Th. 1.9].

Ce problème est celui du convoyeur : si nous devons transporter du sable dans un trou, nous pourrions faire appel à une entreprise de transport (le convoyeur), qui, au lieu de facturer le coût du transport, nous facturerait  $\Phi(x)$  le prix du chargement d'une unité de sable au point  $x$  et  $\Psi(y)$  le prix du déchargement d'une unité de sable au point  $y$ . La contrainte pour le convoyeur est que  $\Phi(x) + \Psi(y) \leq c(x, y)$ , sinon nous irions voir ailleurs car le coût serait trop élevé. Le convoyeur veut charger/décharger au tarif le plus élevé possible. C'est aussi le problème de la dualité de Kantorovich qui s'écrit :

$$\begin{aligned} & \underset{\Phi(x), \Psi(y)}{\text{maximiser}} && \int_{\Omega} \Phi(x) \rho_0(x) dx + \int_{\Omega} \Psi(y) \rho_1(y) dy \\ & \text{sous la contrainte :} && \Phi(x) + \Psi(y) \leq \frac{|x - y|^2}{2}. \end{aligned} \tag{DK}$$

**Définition 2.3.6 (Potentiel de Kantorovich)**

Le  $\Phi$  optimal du problème dual (DK) s'appelle le potentiel de Kantorovich entre  $\rho_0$  et  $\rho_1$ . Le  $\Psi$  optimal est lié à  $\Phi$  par la formule

$$\Psi(y) = \frac{|y|^2}{2} + \left( \Phi - \frac{|\cdot|^2}{2} \right)^*$$

avec  $*$  dénotant la transformée de Legendre. De plus,  $\Psi$  est le potentiel de Kantorovich du transport entre  $\rho_1$  et  $\rho_0$ .

Le théorème suivant montre deux choses, d'abord que le problème dual (DK) redonne aussi la distance de Wasserstein, et que le potentiel de Kantorovich est au cœur de toutes les définitions de la distance de Wasserstein.

**Théorème 2.3.7**

On note par  $C_{DK}$  la valeur du problème (DK),  $C_K$  la valeur du problème (K),  $C_M$  la valeur du problème de Monge et  $C_{BB}$  la valeur du problème (BB).

1. Les problèmes duaux (DK) et (K) sont équivalents, on a :

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = C_{DK} = C_K.$$

2. Il y a également équivalence avec le problème de Monge,  $C_{DK} = C_M$ , et le transport optimal  $T$  est lié au potentiel de Kantorovich  $\Phi$  par :

$$T(x) = \nabla \left( \frac{|x|^2}{2} - \Phi \right). \quad (2.3.11)$$

3. Enfin, il y a équivalence avec le problème de Benamou-Brenier,  $C_{DK} = C_{BB}$ , et la vitesse  $v$  optimale pour le problème de Benamou-Brenier résout l'équation de Burgers de condition initiale :

$$v(0, x) = -\nabla \Phi(x). \quad (2.3.12)$$

La démonstration permet de voir quelques manipulations sur les différents éléments du transport optimal, permettant de passer entre les différentes définitions du transport optimal déjà vues. Elle n'est pas essentielle à la compréhension du reste du manuscrit.

**Démonstration :** La dualité de Kantorovich qui indique que  $C_{DK} = C_K$  est admise, on peut retrouver le résultat et la démonstration dans [Villani, 2003, Th. 1.3.].

Par les équivalences entre les écritures de la distance de Wasserstein qui font l'objet des Théorèmes 2.3.5 et 2.3.2, on a que  $C_{DK} = C_M = C_{BB}$ .

Soit  $(\rho(t, x), v(t, x) = \nabla \phi(t, x))$  le chemin continu optimal pour la formulation Benamou-Brenier. En utilisant les équations de  $\rho$  et  $v$  et  $\phi$ , (voir Remarque 2.3.4) qui sont

$$\partial_t \phi + \frac{|\nabla \phi|^2}{2} = 0 \text{ avec } \phi(0, x) = \Psi(x) - \frac{|x|^2}{2}, \quad (2.3.13a)$$

$$v = \nabla \phi, \quad (2.3.13b)$$

$$\partial_t \rho + \operatorname{div}(\rho \nabla \phi) = 0, \quad (2.3.13c)$$

$$\rho \frac{\partial \phi}{\partial \vec{n}} = 0 \text{ sur } \partial \Omega, \quad (2.3.13d)$$

nous pouvons dans la version Benamou-Brenier du transport optimal remplacer  $v$  par  $\nabla \phi$  puis  $|\nabla \phi|^2$  par  $-2\partial_t \phi$ , puis effectuer des intégrations par parties (I.P.P.) en espace ou en temps, donnant :

$$\begin{aligned} C_{BB} &= \iint_{[0,1] \times \Omega} \frac{1}{2} \rho |\nabla \phi|^2 dt dx \\ &\stackrel{(2.3.13a)}{=} - \iint_{[0,1] \times \Omega} \rho \partial_t \phi dt dx \\ &\stackrel{\text{I.P.P.}}{=} \iint_{[0,1] \times \Omega} \partial_t \rho \phi dt dx - \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx \\ &\stackrel{(2.3.13c)}{=} - \iint_{[0,1] \times \Omega} \operatorname{div}(\rho \nabla \phi) \phi dt dx - \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx \\ &\stackrel{\text{I.P.P.}}{=} \iint_{[0,1] \times \Omega} \rho |\nabla \phi|^2 dt dx - \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx + \underbrace{\oint_{\partial \Omega} \phi \rho \frac{\partial \phi}{\partial \vec{n}} ds}_{=0 \text{ cf. (2.3.13d)}} \\ &= 2C_{BB} - \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx \end{aligned}$$

Il y a égalité entre les premier et dernier termes, donc la valeur de  $C_{BB}$  (qui est aussi la distance de Wasserstein) est :

$$C_{BB} = \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx \quad (2.3.14)$$



Pour faire le lien avec le problème de Kantorovich, on peut utiliser la formule de Hopf-Lax (voir [Evans, 2010, Th. 4 de la partie 3.3.]) qui dit que la solution  $\phi$  de l'équation d'Hamilton-Jacobi (2.3.13a) vérifie

$$\forall x, y \in \Omega, \quad \phi(1, y) - \phi(0, x) \leq \frac{|x - y|^2}{2}.$$

Avec la notation  $\Phi(x) = -\phi(0, x)$  et  $\Psi(y) = \phi(1, y)$ , on a  $\Phi(x) + \Psi(y) \leq \frac{|x-y|^2}{2}$ , on trouve via (2.3.14) que le problème dual est atteint en ce couple  $(\Phi, \Psi)$ . Ainsi,  $\Phi(x) = -\phi(0, x)$  est le potentiel de Kantorovich du transport entre  $\rho_0$  et  $\rho_1$ . De plus, comme  $v = \nabla\phi$ , on a

$$v(0, x) = \nabla\phi(0, x) = -\nabla\Phi(x)$$

et par le Théorème (2.3.2), on a

$$T(x) = \nabla \left( \frac{|x|^2}{2} - \Phi \right). \quad \blacksquare$$

Enfin, avant de passer à la suite, remarquons d'abord que le potentiel de Kantorovich est tangent aux bords du domaine :

### Corollaire 2.3.8 (*Conditions aux bords*)

Le potentiel de Kantorovich  $\Phi$  vérifie

$$\rho_0 \frac{\partial \Phi}{\partial \vec{n}} = 0 \text{ p.p. sur } \partial\Omega.$$

**Démonstration :** C'est une conséquence immédiate de l'équation (2.3.12) et du fait que la vitesse est tangente au bord. \blacksquare

Et puis remarquons que le potentiel de Kantorovich permet de définir une géodésique entre deux densités :

### Corollaire 2.3.9 (*Description du minimiseur à l'aide du potentiel*)

Soit  $\Phi$  le potentiel de Kantorovich du transport entre  $\rho_0$  et  $\rho_1$ . Alors,

$$\rho_1 = (Id - \nabla\Phi)\#\rho_0$$

De plus, le chemin optimal (la géodésique)  $\rho(t)$  reliant  $\rho_0$  et  $\rho_1$ , s'écrit

$$\forall t \in [0, 1], \quad \rho(t) = (Id - t\nabla\Phi)\#\rho_0.$$

**Démonstration :** La première égalité découle directement du théorème 2.3.7. En effet, le transport optimal  $T$  relie  $\rho_0$  et  $\rho_1$  par la contrainte  $\rho_1 = T\#\rho_0$  et  $T$  et  $\Phi$  sont liés par (2.3.11).

La deuxième égalité provient des équations décrivant le couple minimiseur  $(\rho(t), v(t))$  qui ont été détaillées dans le théorème 2.3.2, et qui sont

$$\begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho v) = 0 \\ \frac{\partial v}{\partial t} + v \cdot \nabla v = 0 \\ \rho(0) = \rho_0, \quad v(0) = -\nabla\Phi. \end{cases}$$

Nous pouvons résoudre ces équations par les caractéristiques, comme nous l'avons vu plus tôt dans l'équivalence entre les problèmes de Benamou-Brenier et de Monge. Ces caractéristiques associées à

$v$  s'écrivent

$$T_t(x) = x - t\nabla\Phi$$

et  $\rho$  résout l'équation de continuité avec la vitesse  $v$  donc est donné par

$$\rho(t) = T_t\#\rho_0 = (Id - t\nabla\Phi)\#\rho_0. \quad \blacksquare$$

Enfin, nous présentons une description des fonctions décrites par le symbole  $\#$ , qui découle du précédent corollaire :

**Corollaire 2.3.10**

Soit  $\Phi$  une fonction de l'espace et  $\rho_0 \in \mathcal{P}(\Omega)$ . Soit

$$\rho(t) := (Id - t\nabla\Phi)\#\rho_0.$$

Le  $\rho(t)$  est résout alors le système d'équations :

$$\begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \nabla \phi) = 0 \\ \frac{\partial \phi}{\partial t} + \frac{|\nabla \phi|^2}{2} = 0 \\ \rho(0) = \rho_0, \quad \phi(0, x) = -\Phi(x) \end{cases}$$

### 2.3.4 Résumé

Les quatre différentes versions de l'écriture de la distance de Wasserstein  $\mathcal{W}_2(\rho_0, \rho_1)$  sont résumées dans le tableau 2.1, ainsi que les propriétés des éléments optimisant chacune des versions de la distance de Wasserstein, ainsi que les liens entre ces optima.

On a choisi le potentiel de Kantorovich  $\Phi$  comme élément central des liens entre les différentes versions et différents optima.

## 2.4 L'espace de Wasserstein et son produit scalaire

Dans la suite on va munir  $\mathcal{P}(\Omega)$  de la distance de Wasserstein  $\mathcal{W}_2$  pour calculer la distance entre deux mesures de probabilité. Comme on l'a vu, la distance de Wasserstein a de nombreuses caractéristiques intéressantes. Une dernière caractéristique à étudier et qui va nous servir est la représentation riemannienne de  $(\mathcal{P}(\Omega), \mathcal{W}_2)$ . On parle d'une distance mais on va voir qu'on peut parler plutôt d'un produit scalaire sous-jacent. De même qu'un espace de Hilbert a plus de propriétés qu'un espace de Banach, définir  $(\mathcal{P}(\Omega), \mathcal{W}_2)$  à l'aide d'un produit scalaire donne plus d'informations que de le définir à l'aide simplement de la distance.

Ce produit scalaire, comme on le verra, permettra de définir un gradient de Wasserstein. Le but sera pour une fonction  $\mathcal{J}$  de  $\mathcal{P}(\Omega)$ , d'écrire des variations et notamment le gradient, c'est-à-dire

$$\mathcal{J}(\rho + \varepsilon\eta) = \mathcal{J}(\rho) + \varepsilon\langle \operatorname{grad}\mathcal{J}(\rho), \eta \rangle + \mathcal{O}(\varepsilon).$$

avec  $\langle \cdot, \cdot \rangle$  le produit scalaire choisi,  $\rho$  un élément de  $\mathcal{P}(\Omega)$  et  $\eta$  tel que  $\rho + \varepsilon\eta \in \mathcal{P}(\Omega)$  pour  $\varepsilon > 0$  assez petit.

Lorsque l'on définit un espace à l'aide d'un produit scalaire, on parle de géométrie riemannienne. La représentation riemannienne de  $\mathcal{P}(\Omega)$  provient de Felix Otto dans [Otto, 2001]. Les définitions

TABLE 2.1 – Tableau résumant les différentes écritures de la distance de Wasserstein entre deux densités  $\rho_0$  et  $\rho_1$ , ainsi que quelques propriétés liant les éléments optimaux.

Valeur de $\mathcal{W}_2(\rho_0, \rho_1)^2$	Remarques
<p><b>Version duale</b></p> $\max_{\Phi(x)+\Psi(y)\leq\frac{ x-y ^2}{2}} \int_{\Omega} \Phi(x)\rho_0(x) + \int_{\Omega} \Psi(y)\rho_1(y)$	<p>Le <math>\Phi</math> optimal s'appelle le potentiel de Kantorovich. On a</p> $\rho_0 \frac{\partial \Phi}{\partial \vec{n}} = 0 \text{ p.p. sur } \partial\Omega.$
<p><b>Version de Monge</b></p> <p>minimum <math>T</math> <math>\int_{\Omega}  x - T(x) ^2 \rho_0(x) dx</math>  sous la contrainte : <math>T\#\rho_0 = \rho_1</math> (voir (2.3.2)).</p>	<p>Le transport optimal <math>T</math> s'écrit</p> $T(x) = x - \nabla\Phi(x)$ <p>avec <math>\Phi</math> le potentiel de Kantorovich. Le théorème de Brenier nous dit que <math>x \mapsto \frac{ x ^2}{2} - \Phi(x)</math> est une fonction convexe.</p>
<p><b>Version de Kantorovich</b></p> $\min_{\pi \in \Pi(\rho_0, \rho_1)} \iint_{\Omega \times \Omega}  x - y ^2 d\pi(x, y)$ <p>avec</p> $\Pi(\rho_0, \rho_1) = \left\{ \pi(x, y) \geq 0 : \iint_{\Omega \times \Omega} (\Phi(x) + \Psi(y)) d\pi(x, y) = \int_{\Omega} [\Phi\rho_0 + \Psi\rho_1] \right\}.$	<p>Le support du <math>\pi</math> optimal est concentré le long d'une courbe (<math>y = T(x)</math>) avec <math>T</math> le plan de transport optimal du problème de Monge.</p>
<p><b>Version continue (Benamou-Brenier)</b></p> $\min_{(\rho, v) \in C(\rho_0, \rho_1)} \frac{1}{2} \iint \rho  v ^2 dx dt$ <p>avec</p> $C(\rho_0, \rho_1) := \left\{ \begin{array}{l} (\rho, v) \text{ tel que} \\ \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \rho(t=0) = \rho_0, \rho(t=1) = \rho_1, \\ \rho v \cdot \vec{n} = 0 \text{ sur } \partial\Omega \end{array} \right\}$	<p>La vitesse optimale s'écrit <math>v(t, x) = \nabla\phi(t, x)</math> avec un <math>\phi</math> qui résout l'équation d'Hamilton-Jacobi</p> $\partial_t \phi + \frac{ \nabla\phi ^2}{2} = 0$ <p>et de condition initiale</p> $\phi(0, x) = -\Phi(x).$ <p><math>\rho(t, x)</math> vérifie lui</p> $\rho = (Id - t\nabla\Phi)\#\rho_0.$

données par F. Otto et plus tard dans [Ambrosio et al., 2005] et [Carrillo et al., 2006] ne sont que formelles, c'est-à-dire qu'elles semblent être justes sans que l'on puisse proprement dire ou prouver que  $(\mathcal{P}(\Omega), \mathcal{W}_2)$  est vraiment un espace riemannien. Ainsi, cette représentation riemannienne n'est qu'un outil : on sait qu'on dispose d'un produit scalaire, qu'il fonctionne à peu près (nous donnant un gradient par exemple), nous donnant quelques résultats mais qu'il faudra démontrer rigoureusement en utilisant d'autres outils.

Un produit scalaire est une forme bilinéaire symétrique, défini sur l'espace tangent et donne une métrique sur les vecteurs tangents.

Soit  $\rho(t)$  une courbe dérivable sur  $(\mathcal{P}(\Omega), \mathcal{W}_2)$ , telle que  $\rho(0) = \rho_0$ . Alors on dit que la dérivée  $\rho'(0)$  est un vecteur tangent à  $\mathcal{P}(\Omega)$  en  $\rho_0$ . L'ensemble des vecteurs tangents en  $\rho_0$  est appelé l'espace tangent,  $T_{\rho_0}\mathcal{P}$ . Pour expliciter  $T_{\rho_0}\mathcal{P}$ , on explicite les dérivées pour la distance de Wasserstein. Soit  $\rho(t)$  tel que  $\rho(0) = \rho_0$ , et dérivable en 0 pour la distance de Wasserstein, c'est-à-dire que  $\rho'(0)$  est tel que

$$\mathcal{W}_2(\rho_0 + dt\rho'(0), \rho_0) = \mathcal{O}(dt).$$

On peut montrer ([Ambrosio et al., 2005, Th.8.3.1]) que  $\rho'(0)$  est doté d'une vitesse à potentiel, c'est-à-dire qu'il existe  $\Phi$  tel que  $\rho'(0) = -\text{div}(\rho_0\nabla\Phi)$ . Ainsi on définit l'espace tangent  $T_{\rho_0}\mathcal{P}$  par (cf. [Ambrosio et al., 2005, Def. 8.4.1]) :

$$T_{\rho_0}\mathcal{P} := \left\{ -\text{div}(\rho_0\nabla\Phi), \quad \rho_0 \frac{\partial\Phi}{\partial\vec{n}} = 0 \text{ sur } \partial\Omega \right\}.$$

#### Définition 2.4.1 (*Produit scalaire de Wasserstein*)

Soit  $\rho \in \mathcal{P}(\Omega)$ . Le produit scalaire est défini sur l'espace tangent  $T_\rho\mathcal{P}$  pour deux éléments  $\eta$  et  $\eta'$  de  $T_\rho\mathcal{P}$ , avec  $\eta = -\text{div}(\rho\nabla\Phi)$  et  $\eta' = -\text{div}(\rho\nabla\Phi')$ , par

$$\langle \eta, \eta' \rangle_W = \int_{\Omega} \rho \nabla\Phi \cdot \nabla\Phi'.$$

#### Remarque 2.4.2

Pour bien voir le lien entre ce produit scalaire et la distance de Wasserstein, remarquons que le carré de la norme d'un vecteur tangent  $\eta = -\text{div}(\rho\nabla\Phi)$  vaut

$$\|\eta\|^2 = \langle \eta, \eta \rangle_W = \int_{\Omega} \rho |\nabla\Phi|^2.$$

Si l'on voit  $\eta$  comme une petite perturbation de  $\rho$ , et  $\nabla\Phi$  la vitesse associée à cette perturbation, alors le carré de la norme de  $\eta$  est l'énergie cinétique de cette perturbation. La distance de Wasserstein était définie comme étant le chemin minimisant l'énergie cinétique de transport.

## 2.5 Différentielle de la distance de Wasserstein

Nous allons écrire des fonctions  $\mathcal{J}$  dépendant de  $\rho \in \mathcal{P}(\Omega)$  à l'aide de la distance de Wasserstein. Pour pouvoir minimiser cette fonction, nous aurons besoin de calculer son gradient, qui est l'élément  $\text{grad}\mathcal{J}(\rho)$  dépendant de  $\rho$  tel que

$$\forall \eta \in T_\rho\mathcal{P}, \mathcal{J}(\rho + \varepsilon\eta) = \mathcal{J}(\rho) + \varepsilon\langle \eta, \text{grad}\mathcal{J}(\rho) \rangle + \mathcal{O}(\varepsilon).$$

Pour faire ceci, nous aurons donc besoin de différentier la distance de Wasserstein, c'est à dire calculer au premier ordre  $\mathcal{W}_2(\rho_0 + \varepsilon\eta, \rho_1)$ . C'est ce que nous étudions ici. Soit la fonction suivante

$$\mathcal{F}(\rho_0) = \mathcal{W}_2(\rho_0, \rho_1)^2.$$

**Théorème 2.5.1**

Soit  $\eta \in T_{\rho_0} \mathcal{P}(\Omega)$ ,  $\eta = -\operatorname{div}(\rho_0 \nabla \Phi)$ . Alors,

$$\begin{aligned} \mathcal{F}(\rho_0 + \varepsilon\eta) &= \mathcal{F}(\rho_0) + \varepsilon \langle \eta, \Phi \rangle_2 + \mathcal{O}(\varepsilon) \\ &= \mathcal{F}(\rho_0) + \varepsilon \langle \eta, -\operatorname{div}(\rho_0 \nabla \Phi) \rangle_W + \mathcal{O}(\varepsilon) \end{aligned}$$

avec  $\Phi$  le potentiel de Kantorovich du transport entre  $\rho_0$  et  $\rho_1$ .

La démonstration de ce théorème a été donnée par N. Gigli et repose sur la formulation duale du problème de Kantorovich (DK). De même que pour les autres démonstrations, sa lecture n'est pas indispensable à la compréhension de la suite de ce travail.

**Démonstration :** On suppose que  $\mathcal{F}$  est dérivable. Soit  $\Phi$  et  $\Psi$  les potentiels optimaux, de sorte que

$$\Phi(x) + \Psi(y) \leq \frac{|x - y|^2}{2} \text{ et}$$

$$\mathcal{F}(\rho_0) = \mathcal{W}_2(\rho_0, \rho_1)^2 = \int_{\Omega} \Phi(x) \rho_0(x) dx + \int_{\Omega} \Psi(y) \rho_1(y) dy$$

On a vu notamment que  $\Phi$  est le potentiel de Kantorovich. Ce couple optimal  $(\Phi, \Psi)$  est toujours admissible pour la formulation duale du transport optimal entre  $\rho_0 + \varepsilon\eta$  et  $\rho_1$ , puisque

$$\begin{aligned} \mathcal{F}(\rho_0 + \varepsilon\eta) &= \mathcal{W}_2(\rho_0 + \varepsilon\eta, \rho_1)^2 = \sup_{\tilde{\Phi}(x) + \tilde{\Psi}(y) \leq |x - y|^2} \int_{\Omega} \tilde{\Phi}(x) (\rho_0(x) + \varepsilon\eta) dx + \int_{\Omega} \tilde{\Psi}(y) \rho_1(y) dy \\ &\geq \int_{\Omega} \Phi(x) (\rho_0(x) + \varepsilon\eta) dx + \int_{\Omega} \Psi(y) \rho_1(y) dy \\ &= \mathcal{F}(\rho_0) + \varepsilon \int_{\Omega} \Phi(x) \eta(x) dx \end{aligned}$$

Si bien que pour tout  $\varepsilon > 0$ ,

$$\frac{\mathcal{F}(\rho_0 + \varepsilon\eta) - \mathcal{F}(\rho_0)}{\varepsilon} \geq \int_{\Omega} \Phi(x) \eta(x) dx$$

Et donc en faisant tendre  $\varepsilon$  vers  $0^+$  (dans l'hypothèse que la dérivée existe),

$$D\mathcal{F}(\rho_0) \cdot \eta \geq \int_{\Omega} \Phi(x) \eta(x) dx.$$

On tombe sur l'inégalité inverse en refaisant exactement le même processus pour  $\rho_0 - \varepsilon\eta$ . En effet, on a

$$\begin{aligned} \mathcal{F}(\rho_0 - \varepsilon\eta) &\geq \int_{\Omega} \Phi(x) (\rho_0(x) - \varepsilon\eta) dx + \int_{\Omega} \Psi(y) \rho_1(y) dy \\ &= \mathcal{F}(\rho_0) - \varepsilon \int_{\Omega} \Phi(x) \eta(x) dx. \end{aligned}$$

Et alors, on peut isoler  $(\mathcal{F}(\rho) - \mathcal{F}(\rho_0 - \varepsilon\eta))/\varepsilon$  et faire tendre  $\varepsilon$  vers  $0^+$ , ce qui donne

$$D\mathcal{F}(\rho_0) \cdot \eta \leq \int_{\Omega} \Phi(x) \eta(x) dx$$

et donc  $D\mathcal{F}(\rho_0).\eta = \langle \Phi, \eta \rangle_2$ . Ceci prouve la première ligne de l'égalité.

Pour la deuxième ligne, soit  $\Psi$  tel que  $\eta = -\operatorname{div}(\rho_0 \nabla \Psi)$ . Par intégration par parties, on trouve :

$$\begin{aligned} D\mathcal{F}(\rho_0).\eta &= \int_{\Omega} \Phi(x) \eta(x) \, dx = - \int_{\Omega} \Phi(x) \operatorname{div}(\rho_0 \nabla \Psi) \, dx \\ &= \int_{\Omega} \rho_0 \nabla \Psi \cdot \nabla \Phi \, dx \end{aligned}$$

et comme  $\nabla \Phi$  est tangent au bord (voir le Corollaire 2.3.8), on a :

$$D\mathcal{F}(\rho_0).\eta = \langle \eta, -\operatorname{div}(\rho_0 \nabla \Phi) \rangle_{\rho_0}.$$

Remarque : si  $\nabla \Phi$  n'était pas tangent au bord, alors  $-\operatorname{div}(\rho_0 \nabla \Phi) \notin T_{\rho_0} \mathcal{P}$  et le membre de droite ne serait pas défini. ■

## 2.6 Résolution numérique du problème de transport optimal

Nous réaliserons des expériences numériques avec la distance de Wasserstein, il faudra pouvoir la calculer, ce qui n'est pas trivial à faire a priori. Seul le problème en dimension un dispose d'une solution explicite, voir l'Annexe A.1.

Les premiers algorithmes écrits résolvent des problèmes de combinatoire, on peut citer la méthode hongroise où les densités sont discrétisées en Dirac et une correspondance entre les Diracs initiaux et finaux est cherchée. Cette première méthode est assez longue puisque la complexité est en  $\mathcal{O}(n^4)$ ,  $n$  étant le nombre de Diracs. Depuis, d'autres algorithmes très performants lui ressemblant ont été écrits, comme celui décrit dans [Mérigot, 2011].

Citons aussi des algorithmes de résolution faciles à implémenter et rapide basée sur une régularisation des contraintes de la distance de Wasserstein. Cela s'appelle la régularisation d'entropie, décrite dans l'Annexe A.6.

D'autres types d'algorithmes ont été écrits et résolvent la formulation continue de Benamou-Brenier

$$\inf_{(\rho, v) \in C(\rho_0, \rho_1)} \iint_{[0,1] \times \Omega} \rho |v|^2 \, dt dx$$

puisqu'après un changement de variable  $m = \rho v$ , le problème est convexe et les contraintes sont affines. Ainsi, on peut faire appel à une grande classe d'algorithmes pour résoudre ce problème, de l'algorithme de descente du gradient (qui n'est pas très efficace) à des lagrangiens augmentés, en passant par les algorithmes type Primal-Dual. Quelques-uns de ces algorithmes sont présentés en Annexe A. Ce sont des algorithmes itératifs assez rapides, qui incrémentent progressivement les valeurs des différentes variables impliquées. L'avantage de ces algorithmes est qu'ils sont très souples : on peut ajouter des termes supplémentaires très facilement à la fonction coût, comme par exemple dans [Hug et al., 2015, Bouharguane et al., 2012], ou pour calculer la distance de Wasserstein-Fisher-Rao (cf. la section suivante) qui est très intéressante pour calculer la distance de Wasserstein entre deux densités de masses différentes.

Pensons aussi au fait que la convergence d'un algorithme itératif dépend de sa condition initiale : plus on part loin de la solution, plus le nombre d'itérations est grand. Or, si l'on a à calculer un grand nombre de distances  $\mathcal{W}_2(\rho_0^n, \rho_1^n)$ , comme on sera amené à le faire, on peut partir de la solution obtenue au calcul de  $\mathcal{W}_2(\rho_0^n, \rho_1^n)$  pour initialiser le calcul de  $\mathcal{W}_2(\rho_0^{n+1}, \rho_1^{n+1})$ , ce qui réduit fortement le temps de calcul.

Dans la suite nous utiliserons la formule exacte, en une dimension, car plus rapide et on évite les erreurs dues aux approximations. Nous essaierons tout de même l'algorithme itératif (donc plus long et inexact) Primal-Dual et l'approximation par régularisation d'entropie, en vue d'applications en dimensions supérieures.

## 2.7 Transport optimal pour des densités de masses différentes

Nous avons vu que la distance de Wasserstein n'était définie que pour des densités de même masse, ce qui limite un peu ses applications. Cette section traite du transport entre des densités de masses différentes. En effet, on retrouve dans la littérature quelques initiatives traitant ce point-ci.

Cette section n'est pas indispensable pour la suite du manuscrit, étant donné que nous utiliserons principalement la distance de Wasserstein  $\mathcal{W}_2$ . L'idée de cette section est de présenter la distance  $WF$ , de montrer qu'elle se comporte de manière semblable à la distance de Wasserstein, et d'énoncer quelques résultats prouvant que ce que l'on fera avec  $\mathcal{W}_2$  (contraignant à utiliser des densités de masses égales) pourra être aussi fait avec  $WF$ .

### 2.7.1 Premières initiatives de transport optimal de densités de masses différentes

Pour résoudre ce problème, dans [Piccoli et Rossi, 2014] les auteurs proposent, sur une idée originale de Benamou [Benamou, 2003], de relaxer la condition d'égalité des intégrales en ajoutant deux variables en plus. Ainsi, pour  $a > 0$  et  $b > 0$  deux constantes, on définit

$$\mathcal{W}_2^{a,b}(\rho_0, \rho_1)^2 := \inf_{\substack{\tilde{\rho}_0 \geq 0, \tilde{\rho}_1 \geq 0 \\ \int_{\Omega} [\tilde{\rho}_0 - \tilde{\rho}_1] dx = 0}} a^2 (\|\rho_0 - \tilde{\rho}_0\|_1^2 + \|\rho_1 - \tilde{\rho}_1\|_1^2) + b^2 \mathcal{W}_2^2(\tilde{\rho}_0, \tilde{\rho}_1)$$

Les auteurs montrent que c'est une distance. De plus, la version Benamou-Brenier de cette distance est

$$\mathcal{W}_2^{a,b}(\rho_0, \rho_1)^2 := \inf_{(\rho, v, S) \in C(\rho_0, \rho_1)} a^2 \int_0^1 \|S(t, \cdot)\|_1^2 dt + b^2 \iint_{[0,1] \times \Omega} \rho |v|^2 dt dx \quad (2.7.1)$$

avec  $C = \{(\rho, v, S) : \partial_t \rho + \operatorname{div}(\rho v) = S\}$ . On a donc un terme source dans l'équation de transport, qui traduit l'apparition et la disparition de matière. Ce terme source est pénalisé.

Cependant, comme le fait remarquer [Chizat et al., 2015b], pénaliser ainsi le terme source n'est pas parfait. En effet, les géodésiques associées à la distance  $\mathcal{W}_2^{a,b}$  ont toujours un comportement double : une partie de la masse est transportée, à cause du terme  $b^2 \mathcal{W}_2$  dans la définition de la distance, tandis qu'une autre partie de la masse apparaît/disparaît, tel une géodésique euclidienne, et ce à cause du terme source. Ce comportement n'est pas vraiment souhaitable car cela fait très artificiel de voir les deux effets simultanément.

### 2.7.2 Distance de Wasserstein-Fisher-Rao

Les auteurs de [Chizat et al., 2015b] propose une distance qui permettrait d'avoir soit un comportement de transport, soit un comportement d'apparition/disparition. Ceci est obtenu en pénalisant le terme source différemment, à l'aide d'une métrique dite de Fisher-Rao. La distance est obtenue est alors nommée distance de Wasserstein-Fisher-Rao, que l'on écrira  $WF$ , et se présente comme une interpolation entre la distance de Wasserstein et celle de Fisher-Rao. Elle se définit différemment de (2.7.1) à cause du poids sur le terme source :

$$WF(\rho_0, \rho_1)^2 = \inf_{(\rho, v, g) \in C(\rho_0, \rho_1)} \frac{1}{2} \iint_{[0,1] \times \Omega} [\rho |v|^2 + \delta \rho g^2] dt dx \quad (2.7.2)$$

où l'ensemble  $C$  est l'ensemble des chemins décrits par une équation de continuité avec terme source pouvant combler le défaut de masse,

$$C(\rho_0, \rho_1) := \{(\rho, v, g) : \partial_t \rho + \operatorname{div}(\rho v) = \rho g, \rho(0) = \rho_0, \rho(1) = \rho_1\}.$$

Le terme  $\delta$  est le poids sur le terme de Fisher-Rao de la distance  $WF$ . Avec ce choix de pénalisation du terme source, les géodésiques ont un comportement unique : soit la matière se transporte en changeant éventuellement de volume, soit la matière apparaît/disparaît, mais rarement les deux à la fois. Ces comportements dépendent de la valeur de  $\delta$ , comme nous allons le voir.

L'article [Chizat et al., 2015b] expose l'exemple du transport optimal entre deux Diracs  $\rho_0$  et  $\rho_1$  respectivement placés aux points  $x_0$  et  $x_1$ . Les deux densités ne sont pas nécessairement de la même masse. Le comportement de la géodésique entre ces deux Diracs dépend de la valeur de  $\delta$  et de la valeur seuil  $\delta_0 = \frac{1}{\pi}|x_0 - x_1|$ , indépendant de la valeur des masses.

- Si  $\delta > \delta_0$ , c'est-à-dire dans le cas où le poids du terme source dans la fonction coût est fort, la géodésique est un Dirac qui se transporte comme avec la distance de Wasserstein. Cependant, même si les masses initiales et finales sont égales, la masse sur la géodésique peut varier.
- Si  $\delta < \delta_0$ , lorsque le poids sur le terme source est faible, le comportement est plutôt à l'avantage de ce terme source : la géodésique est alors simplement la même qu'avec la distance euclidienne.
- Dans le cas limite  $\delta = \delta_0$ , il y a une infinité de géodésiques. Ce cas ne devrait pas se produire chez nous puisqu'on ne traite pas des Diracs mais des mesures à densités.

Ainsi un faible  $\delta$  permet d'éviter de grands déplacements de masses. On prend l'exemple suivant. La densité  $\rho_0$  est une masse localisée, tandis que  $\rho_1$  composé d'une partie proche de  $\rho_0$  et d'une autre plus éloignée. On trace sur la figure 2.13 ces deux densités ainsi que leur moyenne de Wasserstein qui consiste donc en deux parties qui se déplacent.

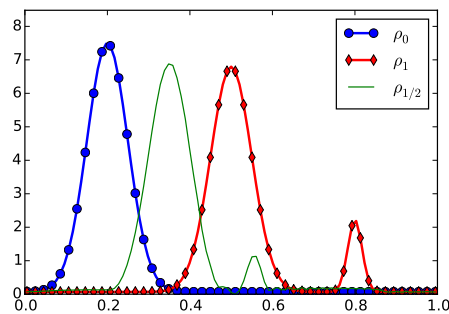


FIGURE 2.13 – On affiche ici la moyenne de Wasserstein de deux densités, l'une étant d'un seul bloc alors que l'autre est en deux parties. La conséquence est que la moyenne de Wasserstein est en deux parties.

Sur la figure 2.14, on affiche la moyenne pour la distance  $WF$  pour différentes valeurs de  $\delta$  ( $\delta = 0,5, 0,1$  et  $0,01$ ), en fonction de l'écartement des différentes parties de  $\rho_0$  et  $\rho_1$ , sachant qu'il y a un écart de  $0,3$  entre  $\rho_0$  et la grosse partie de  $\rho_1$ , et un écart de  $0,6$  avec la petite partie de  $\rho_1$ . Ce qui correspond à des valeurs limites de  $\delta$  de  $0,1$  et de  $0,2$  à peu près. Pour le premier  $\delta = 0,5$ , on constate un transport classique ; pour le deuxième  $\delta = 0,1$ , la petite masse a disparu, remplacée par une apparition type euclidien ; et pour le dernier  $\delta = 0,01$  la moyenne ressemble à une moyenne  $\mathcal{L}^2$ .



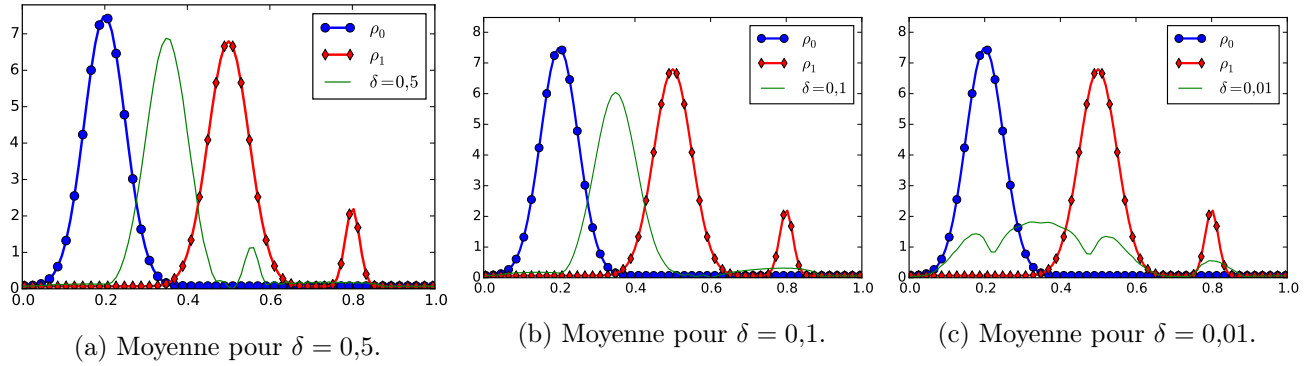


FIGURE 2.14 – Différentes moyennes  $WF$  de  $\rho_0$  et  $\rho_1$  pour différentes valeurs de  $\delta$ . À comparer avec la figure 2.13 où est tracée la moyenne de Wasserstein.

**Influence de la taille des masses.** On a dit que la distance  $WF$  pouvait calculer une géodésique entre deux densités de masses différentes. On affiche sur la figure 2.15 la valeur de la moyenne pour d’abord deux densités de même masse, et ensuite dans le cas où l’une des deux a été multipliée par 2. On constate que la moyenne est la même à un agrandissement près, et aussi à léger déplacement près.

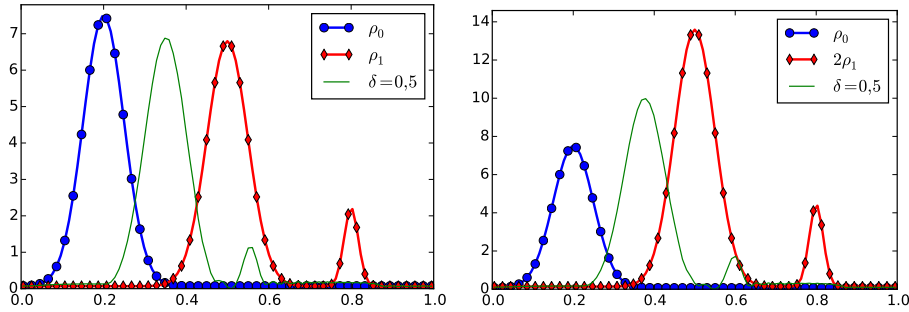


FIGURE 2.15 – Moyenne  $WF$  pour des densités de même masse à gauche (les mêmes  $\rho_0$  et  $\rho_1$  que dans la figure 2.13), et de masses différentes à droite, pour  $\delta = 0,5$ .

### Remarque 2.7.1 (Implémentation numérique)

Pour fournir ces résultats, nous avons implémenté un algorithme de type Primal-Dual résolvant la formulation (2.7.2) de la distance  $WF$ . Cet algorithme est similaire à celui du calcul de la distance de Wasserstein. Nous le présentons en annexe A.7.

### 2.7.3 L’espace Riemannien associé à la distance Wasserstein-Fisher-Rao

Tout comme nous l’avons fait pour la distance de Wasserstein, nous allons rapidement reprendre les éléments de la différentiabilité de la distance Fisher-Rao. D’abord nous allons voir l’équation des géodésiques, ensuite le produit scalaire naturel associé à cette distance, et enfin sa différentielle. Ces éléments ont été établis durant cette thèse par simple manipulation des conditions d’optimalité.

Nous noterons par  $P(\Omega)$  l’espace des fonctions simplement positives de  $\Omega$ .

**L’équation des géodésiques.** Pour la distance de Wasserstein, nous avons trouvé l’équation des géodésiques en comparant les problèmes de Monge et la formulation Benamou-Brenier. Ici, elle

peut être trouvée en écrivant les contraintes de la définition (2.7.2) à l'aide d'un multiplicateur de Lagrange. Le problème se réécrit :

$$\inf_{\rho, v, g} \sup_{\phi} \mathcal{L}(\rho, v, g, \phi)$$

avec

$$\mathcal{L}(\rho, v, g, \phi) = \iint_{\Omega \times [0,1]} \left[ \frac{1}{2} \rho |v|^2 + \frac{\delta}{2} \rho g^2 - \partial_t \phi \rho - \rho v \cdot \nabla \phi - \rho g \phi \right] dx dt + \int_{\Omega} [\rho_1 \phi(1) - \rho_0 \phi(0)] dx.$$

Les conditions d'Euler-Lagrange donnent l'équation du point-selle et de la géodésique :

**Proposition 2.7.2**

La géodésique  $\rho(t)$  entre  $\rho_0$  et  $\rho_1$  résout le système d'équations :

$$\begin{cases} \frac{\partial \phi}{\partial t} + \frac{|\nabla \phi|^2}{2} + \frac{1}{\delta} \frac{\phi^2}{2} = 0 \\ \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \nabla \phi) = \frac{1}{\delta} \rho \phi. \end{cases} \quad (2.7.3)$$

La vitesse  $v$  vaut  $v = \nabla \phi$  et la source  $g$  vaut  $g = \frac{1}{\delta} \phi$ . Les conditions initiales sont  $\rho(0) = \rho_0$ , et  $\Phi(0)$  est de telle sorte que  $\rho(1) = \rho_1$ . Enfin, la distance  $WF$  se réécrit alors :

$$WF(\rho_0, \rho_1)^2 = \frac{1}{2} \iint_{\Omega \times [0,1]} \rho [|\nabla \phi|^2 + \frac{1}{\delta} \phi^2] dt dx. \quad (2.7.4)$$

**Le produit scalaire associé.** La formulation des géodésiques (de  $\partial_t \rho$  notamment) donne le produit scalaire suivant :

**Définition 2.7.3**

Soit  $\rho_0 \in P(\Omega)$ . L'espace tangent  $T_{\rho_0} P$  est l'ensemble des  $\eta$  décrits par un potentiel  $\Phi$  tel que

$$\eta = -\operatorname{div}(\rho_0 \nabla \Phi) + \frac{1}{\delta} \rho_0 \Phi$$

avec la condition de Neumann homogène au bord

$$\rho_0 \frac{\partial \Phi}{\partial \bar{n}} = 0 \text{ sur } \partial \Omega.$$

Alors, pour  $\eta$  et  $\eta'$  de potentiels respectifs  $\Phi$  et  $\Phi'$ , on définit le produit scalaire :

$$\langle \eta, \eta' \rangle_{WF} := \int_{\Omega} \left[ \rho_0 \nabla \Phi \cdot \nabla \Phi' + \frac{1}{\delta} \rho_0 \Phi \Phi' \right] dx.$$

On peut voir que ce produit scalaire est associé à la distance  $WF$ . En effet, la norme de  $\eta$  est

$$\|\eta\|^2 = \int_{\Omega} \left[ \rho_0 |\nabla \Phi|^2 + \frac{1}{\delta} \rho_0 \Phi^2 \right] dt dx$$

qui évoque la distance  $WF$  en remarquant que  $\Phi = \delta g$ .

**La différentielle** Nous calculons la différentielle de la distance  $WF$ , dont nous aurons besoin pour minimiser des fonctions écrites à l'aide de cette distance. La différentielle n'est pas un élément qui a été trouvé dans la littérature, aussi en faisons-nous une preuve "maison".

**Théorème 2.7.4 (Différentielle de  $WF$ )**

Soit  $\rho_0$  et  $\rho_1$  deux fonctions positives, soit  $\phi(t, x)$  le potentiel qui décrit la géodésique entre  $\rho_0$  et  $\rho_1$  comme dans (2.7.3). On pose  $\Phi(x) := -\phi(0, x)$ . Soit  $\eta_0$  tel que  $\rho_0 + \varepsilon\eta_0 \geq 0$  pour  $\varepsilon > 0$  suffisamment petit ; on a :

$$\begin{aligned} WF(\rho_0 + \varepsilon\eta_0, \rho_1)^2 &= WF(\rho_0, \rho_1)^2 + \varepsilon \int_{\Omega} \eta(x)\Phi(x) dx + \mathcal{O}(\varepsilon) \\ &= WF(\rho_0, \rho_1)^2 + \varepsilon \left\langle \eta, -\operatorname{div}(\rho_0 \nabla \Phi) + \frac{1}{\delta} \rho_0 \Phi \right\rangle_{WF} + \mathcal{O}(\varepsilon) \end{aligned}$$

**Démonstration :** Il y a plusieurs moyens de prouver cette différentielle : soit en utilisant la dualité de Kantorovich formulée pour la distance  $WF$  qu'on peut trouver dans [Chizat et al., 2015a, Cor.5.6.]. Mais il faudrait prouver en plus que le potentiel de Kantorovich  $\Phi(x)$  vaut l'opposée de la condition initiale  $-\phi(0, x)$ , ce que nous n'avons fait. Nous pouvons aussi utiliser l'équation des géodésiques.

Soit  $\eta_0$  une perturbation de  $\rho_0$ , et soit  $(\rho^\varepsilon, \phi^\varepsilon)$  décrivant la géodésique (2.7.3) entre  $\rho_0 + \varepsilon\eta_0$  et  $\rho_1$ . Nous supposons que nous pouvons linéariser cette géodésique en  $\varepsilon = 0$  par :

$$\begin{aligned} \rho^\varepsilon &= \rho^0 + \varepsilon\eta + \mathcal{O}(\varepsilon) \\ \phi^\varepsilon &= \phi^0 + \varepsilon\Psi + \mathcal{O}(\varepsilon) \end{aligned}$$

avec  $(\rho^0, \phi^0)$  les variables décrivant la géodésique entre  $\rho_0$  et  $\rho_1$ . Pour trouver les équations sur  $\eta(t, x)$  et  $\Psi(t, x)$ , nous linéarisons les équations (2.7.3) de  $\rho^\varepsilon, \phi^\varepsilon$  en  $\varepsilon = 0$ , donnant :

$$\begin{cases} \frac{\partial \psi}{\partial t} + \nabla \phi^0 \cdot \nabla \psi + \frac{1}{\delta} \phi^0 \psi = 0 \\ \frac{\partial \eta}{\partial t} + \operatorname{div}(\eta \nabla \phi^0) + \operatorname{div}(\rho^0 \nabla \psi) = \frac{1}{\delta} \eta \phi^0 + \frac{1}{\delta} \rho^0 \psi \end{cases} \quad (2.7.5)$$

La distance  $WF$  est elle aussi linéarisée, en utilisant sa formulation à l'aide de  $\rho^\varepsilon$  et  $\phi^\varepsilon$  (cf. (2.7.4)) :

$$\begin{aligned} WF(\rho_0 + \varepsilon\eta_0, \rho_1)^2 &= \frac{1}{2} \iint_{[0,1] \times \Omega} \rho^\varepsilon \left( |\nabla \phi^\varepsilon|^2 + \frac{1}{\delta} |\phi^\varepsilon|^2 \right) dt dx \\ &= \frac{1}{2} \iint_{[0,1] \times \Omega} \rho^0 \left( |\nabla \phi^0|^2 + \frac{1}{\delta} |\phi^0|^2 \right) dt dx + \\ &\quad + \varepsilon \iint_{[0,1] \times \Omega} \underbrace{\frac{1}{2} \left[ \eta \left( |\nabla \phi^0|^2 + \frac{1}{\delta} |\phi^0|^2 \right) + \rho^0 \left( \nabla \phi^0 \cdot \nabla \Psi + \frac{1}{\delta} \phi^0 \Psi \right) \right]}_{=: I} dt dx \\ &= WF(\rho_0, \rho_1)^2 + \varepsilon I + \mathcal{O}(\varepsilon) \end{aligned}$$

Le  $I$  est la différentielle. Nous cherchons à l'exprimer comme un produit scalaire avec  $\eta_0$ . Pour ceci nous faisons des intégrations par partie successives, en utilisant les équations (2.7.5) ; Par simplicité

d'écriture, nous écrirons dans ce calcul  $\phi$  pour  $\phi^0$  et  $\rho$  pour  $\rho^0$  :

$$\begin{aligned}
I &= \iint_{[0,1] \times \Omega} \eta \left[ \underbrace{\frac{1}{2}|\nabla\phi|^2 + \frac{1}{2}\frac{\phi^2}{\delta}}_{=-\partial_t\phi} \right] + \underbrace{\rho[\nabla\phi \cdot \nabla\psi + \frac{1}{\delta}\phi\psi]}_{=:A} \\
&= \iint_{[0,1] \times \Omega} \phi \partial_t \eta + \int_{\Omega} [\phi(0)\eta(0) - \phi(1)\eta(1)] + A \\
&= \iint_{[0,1] \times \Omega} \phi \left( -\operatorname{div}(\eta\nabla\phi) - \operatorname{div}(\rho\nabla\psi) + \frac{1}{\delta}\eta\phi + \frac{1}{\delta}\rho\psi \right) + \int_{\Omega} \underbrace{[\phi(0)\eta(0) - \phi(1)\eta(1)]}_{=\phi(0)\eta_0} + A \\
&= \iint_{[0,1] \times \Omega} \eta |\nabla\phi|^2 + \rho\nabla\phi \cdot \nabla\psi + \frac{1}{\delta}\eta\phi^2 + \frac{1}{\delta}\rho\psi\phi + \int_{\Omega} \phi(0)\eta_0 + A \\
&= \iint_{[0,1] \times \Omega} \eta \left[ |\nabla\phi|^2 + \frac{\phi^2}{\delta} \right] + 2A + \int_{\Omega} \phi(0)\eta_0 \\
&= 2I + \int_{\Omega} \phi(0)\eta_0
\end{aligned}$$

Ainsi, comme le premier et dernier termes sont égaux, on a la valeur de  $I$  et :

$$WF(\rho_0 + \varepsilon\eta_0, \rho_1)^2 = WF(\rho_0, \rho_1)^2 - \int_{\Omega} \phi(0)\eta_0 + \mathcal{O}(\varepsilon)$$

ce qui prouve la première égalité du théorème. Pour la deuxième, si  $\eta$  est à potentiel  $\Psi$ , c'est-à-dire si :

$$\eta = -\operatorname{div}(\rho_0\nabla\Psi) + \frac{1}{\delta}\rho_0\Psi$$

avec les conditions aux bords de Neumann homogènes, alors le produit scalaire se transforme par intégration par parties :

$$\begin{aligned}
\left\langle \eta, -\operatorname{div}(\rho_0\nabla\Phi) + \frac{1}{\delta}\rho_0\Phi \right\rangle_{WF} &= \int_{\Omega} \rho_0\nabla\Psi\nabla\Phi + \delta\rho_0\Psi\Phi \\
&= \int_{\Omega} -\operatorname{div}(\rho_0\nabla\Psi)\Phi + \delta\rho_0\Psi\Phi \\
&= \int_{\Omega} \eta\Phi.
\end{aligned}$$

■

## 2.8 La distance $\mathcal{H}^{-1}$

La distance  $\mathcal{H}^{-1}$  présente beaucoup de similarités avec la distance de Wasserstein. Elle est en effet capable de mesurer la distance qui sépare deux objets séparés et n'est pas sensible au bruit. Elle a également l'air convexe par rapport au déplacement, comme nous allons le voir. Elle est même plus simple à calculer. Elle présente cependant un défaut : le barycentre de deux densités pour cette distance est le même que le barycentre pour la distance euclidienne.

Nous l'utiliserons quelques fois dans la suite. Elle peut servir aussi comme outil de compréhension à la distance de Wasserstein, les deux étant semblables en certains points. Mais comme le travail de cette thèse s'est plutôt concentré sur la distance de Wasserstein, offrant de résultats plus probants, cette section pourra être omise dans une première lecture sans gêner la compréhension de la suite du manuscrit.

La distance de Sobolev  $\mathcal{H}^{-1}$  est originellement une distance vivant sur l'espace dual des fonctions  $\mathcal{H}_0^1$ , c'est-à-dire des fonctions dérivables au sens faible, et nulles au bord de  $\Omega$ . Nous changerons légèrement sa définition pour pouvoir l'utiliser sur l'espace  $\mathcal{P}(\Omega)$ .

### 2.8.1 Définition classique de la distance $\mathcal{H}^{-1}$

L'espace  $\mathcal{H}^{-1}(\Omega)$  est le dual de l'espace  $\mathcal{H}_0^1(\Omega)$  qui est l'espace des fonctions dérivables faiblement et nulles au bord de  $\Omega$ . On rappelle que le produit scalaire et la norme sur  $\mathcal{H}_0^1$  sont

$$\langle \Phi, \Psi \rangle_{\mathcal{H}_0^1} = \int_{\Omega} \nabla \Phi \cdot \nabla \Psi \, dx, \quad \|\Phi\|_{\mathcal{H}_0^1}^2 = \int_{\Omega} |\nabla \Phi|^2 \, dx.$$

L'espace  $\mathcal{H}_0^1(\Omega)$  muni de ce produit scalaire est un espace de Hilbert. L'espace dual de  $\mathcal{H}_0^1(\Omega)$ , que l'on nomme  $\mathcal{H}^{-1}$ , est l'espace des fonctions linéaires continues sur  $\mathcal{H}_0^1$ . Pour  $T \in \mathcal{H}^{-1}$  on a la notation suivante de crochet de dualité :

$$T : \Psi \in \mathcal{H}_0^1 \mapsto \langle T, \Psi \rangle_{\mathcal{H}^{-1} \times \mathcal{H}_0^1} \in \mathbb{R}$$

et la norme de  $T$  est définie par

$$\|T\|_{-1} = \sup_{\Psi \in \mathcal{H}_0^1} \frac{\langle T, \Psi \rangle_{\mathcal{H}^{-1} \times \mathcal{H}_0^1}}{\|\Psi\|_{\mathcal{H}_0^1}}.$$

Chaque élément  $T$  de  $\mathcal{H}^{-1}$  étant une forme linéaire continue de  $\mathcal{H}_0^1$ , le théorème de représentation de Riesz nous dit qu'il existe  $\Phi \in \mathcal{H}_0^1$  tel que :

$$\forall \Psi \in \mathcal{H}_0^1, \quad \langle T, \Psi \rangle_{\mathcal{H}^{-1} \times \mathcal{H}_0^1} = \langle \Phi, \Psi \rangle_{\mathcal{H}_0^1} = \int_{\Omega} \nabla \Phi \cdot \nabla \Psi \, dx$$

Premièrement, la norme de  $T$  vaut  $\|T\|_{-1} = \|\Phi\|_{\mathcal{H}_0^1}$ . De plus, en remarquant que l'application  $\Psi \mapsto \int_{\Omega} \nabla \Phi \cdot \nabla \Psi \, dx$  est exactement  $-\Delta \Phi$  au sens faible, on a

$$T = -\Delta \Phi.$$

Ce Laplacien est donc inversible sur  $\mathcal{H}^{-1}$  avec les conditions aux bords de Dirichlet, on parle de Laplacien-Dirichlet. On introduit la notation  $\Phi = -\Delta^{-1}T$ ,

$$\Phi = -\Delta^{-1}f \text{ si } \begin{cases} -\Delta \Phi = f & \text{dans } \Omega \\ \Phi = 0 & \text{sur } \partial\Omega. \end{cases}$$

Avec cette notation, la norme de  $T$  vaut :

$$\|T\|_{-1}^2 = \int_{\Omega} |\nabla \Delta^{-1}T|^2 \, dx$$

Très schématiquement, on peut considérer la norme  $\mathcal{H}^{-1}$  comme la norme des "primitives", tout comme la norme  $\mathcal{H}^1$  est la norme des dérivées ; car  $\nabla \Delta^{-1}$  est vu comme un inverse de l'opérateur divergence et donc intègre les fonctions.

#### Remarque 2.8.1

*L'espace  $\mathcal{H}^{-1}$  est plus grand que  $\mathcal{L}^2$ . On y retrouve toutes les fonctions  $\mathcal{L}^2$ , mais pas seulement. Par exemple, sur  $\Omega = [-1, 1]$ , le Dirac en 0 noté  $\delta_0$  est élément de  $\mathcal{H}^{-1}$  : c'est le Laplacien de la fonction chapeau  $x \mapsto |x| - 1$  qui appartient bien à  $\mathcal{H}_0^1([-1, 1])$ . Cependant, pour des raisons numériques on utilisera uniquement la norme  $\mathcal{H}^{-1}$  sur un espace de fonction plus abordable, tel que  $\mathcal{L}^2$  justement, ou les fonctions continues.*

### 2.8.2 Intérêt : mesurer le déplacement

Soit  $\Omega$  un intervalle de  $\mathbb{R}$ , et considérons deux masses localisées distantes de  $d$ . Par exemple,  $\rho_0$  et  $\rho_1 = \rho_0(x - d)$ . Alors, comme on l'a fait plus tôt dans la figure 2.4 page 40, on trace dans la figure 2.16 en fonction du déplacement  $d$  la distance  $\mathcal{L}^2$ ,  $\mathcal{W}_2$  et  $\mathcal{H}^{-1}$  entre  $\rho_0$  et  $\rho_1 = \rho_0(x - d)$ . On remarque ici que la distance  $\mathcal{H}^{-1}$  est presque convexe par rapport au déplacement.

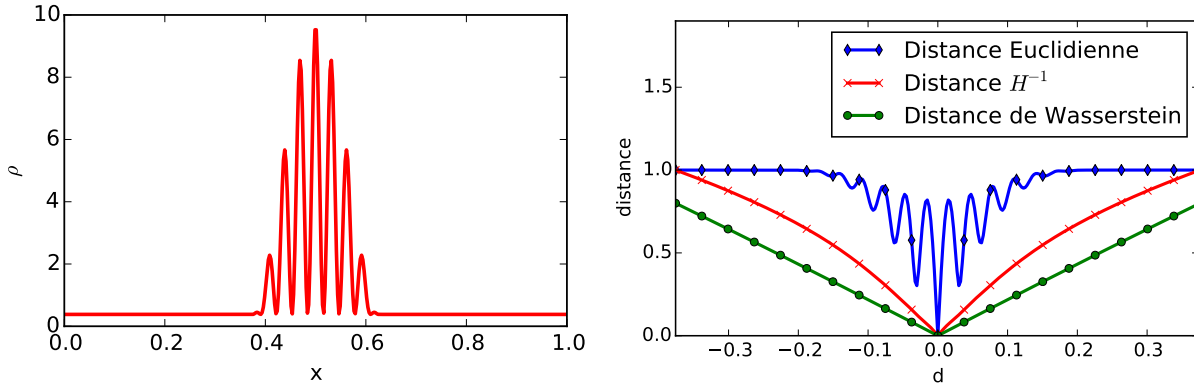


FIGURE 2.16 – À gauche  $\rho_0$ , une fonction oscillante. On pose  $\rho_1 = \rho_0(\cdot - d)$  et on trace à droite, en fonction du déplacement  $d$ , la valeur de  $\text{dist}(\rho_0, \rho_1)$  avec  $\text{dist}$  la distance  $\mathcal{L}^2$ ,  $\mathcal{H}^{-1}$  ou  $\mathcal{W}_2$  (distance normalisée).

### 2.8.3 Les distances $\mathcal{H}^{-1}$ et lien avec la distance de Wasserstein

La distance  $\mathcal{H}^{-1}$  est définie pour deux fonctions  $\mathcal{L}^2$  comme la norme  $\mathcal{H}^{-1}$  de la différence,

$$\mathcal{H}^{-1}(\rho_0, \rho_1) = \int_{\Omega} |\nabla \Delta^{-1}(\rho_0 - \rho_1)|^2 dx$$

avec  $\Delta^{-1}$  la notation pour la résolution du Laplacien-Dirichlet, c'est-à-dire pour  $f \in \mathcal{L}^2$ ,

$$\Phi = -\Delta^{-1}f \text{ si } \begin{cases} -\Delta\Phi = f & \text{dans } \Omega \\ \Phi = 0 & \text{sur } \partial\Omega. \end{cases}$$

En fait on peut très bien définir une autre distance  $\mathcal{H}^{-1}$  en changeant les conditions aux bords. On peut avoir des conditions de Neumann, et le Laplacien-Neumann s'écrit

$$\Phi = -\Delta^{-1}f \text{ si } \begin{cases} -\Delta\Phi = f & \text{dans } \Omega \\ \frac{\partial\Phi}{\partial\vec{n}} = 0 & \text{sur } \partial\Omega \end{cases}$$

qui est bien défini pour  $f \in \mathcal{L}^2$ , et seulement si  $f$  est d'intégrale nulle sur  $\Omega$ . Pour doter  $\mathcal{P}(\Omega)$  d'une métrique  $\mathcal{H}^{-1}$ , on peut plutôt utiliser la définition utilisant ce Laplacien-Neumann-ci, et ce pour trois raisons :

- Tout d'abord, pour  $\rho_0$  et  $\rho_1$  dans  $\mathcal{P}(\Omega)$ , la différence  $\rho_0 - \rho_1$  est justement de moyenne nulle. C'est un signe pour utiliser le Laplacien-Neumann.

- Nous rencontrerons un unique détail qui conditionnera le choix des conditions aux bords. On le verra et nous étudierons dans l'annexe C ce choix. Jusqu'à cette annexe, on supposera que choisir le Laplacien-Neumann sera satisfaisant.
- Le lien avec la distance de Wasserstein. On peut linéariser la distance de Wasserstein autour de 1 avec cette distance,

$$\sqrt{2} \mathcal{W}_2(1 + \varepsilon\eta_0, 1 + \varepsilon\eta_1) = \varepsilon \|\eta_1 - \eta_0\|_{-1} + \mathcal{O}(\varepsilon)$$

avec la distance  $\mathcal{H}^{-1}$  utilisant ce Laplacien-Neumann-ci. L'ébauche de la preuve est donnée ici, et se base sur la partie sur linéarisation de  $\mathcal{W}_2$  présentée dans [Villani, 2003, Sec. 7.6].

**Démonstration :** Soit  $(\rho_\varepsilon, v_\varepsilon = \nabla\Phi_\varepsilon(t, x))$  le chemin optimal correspondant à  $\mathcal{W}_2(1 + \varepsilon\eta_0, 1 + \varepsilon\eta_1)$ . On rappelle que la vitesse est tangente au bord et donc que

$$\frac{\partial\Phi_\varepsilon}{\partial\vec{n}} = 0 \text{ sur } \partial\Omega.$$

On le verra, c'est ce point-ci qui dit qu'il vaut mieux utiliser un Laplacien-Neumann. On linéarise ces  $\rho_\varepsilon$  et  $\Phi_\varepsilon$ , pour ceci on suppose qu'ils sont dérivables en  $\varepsilon$ , et on effectue le développement limité

$$\begin{aligned} \rho_\varepsilon &= 1 + \varepsilon\eta(t, x) + \mathcal{O}(\varepsilon) \\ \Phi_\varepsilon &= \Phi_0 + \varepsilon\Psi(t, x) + \mathcal{O}(\varepsilon). \end{aligned} \tag{2.8.1}$$

Pour trouver les équations sur  $\Psi$  et  $\eta$ , on utilise les équations du transport optimal qui sont

$$\begin{cases} \frac{\partial\rho_\varepsilon}{\partial t} + \operatorname{div}(\rho_\varepsilon\nabla\Phi_\varepsilon) = 0 \\ \frac{\partial\Phi_\varepsilon}{\partial t} + \frac{|\nabla\Phi_\varepsilon|^2}{2} = 0. \end{cases}$$

En linéarisant (c'est-à-dire en remplaçant  $\rho_\varepsilon$  et  $\Phi_\varepsilon$  par leurs développements limités (2.8.1), et en faisant tendre  $\varepsilon \rightarrow 0$ ), on obtient à l'ordre zéro que

$$\Phi_0 = 0$$

et au premier ordre que

$$\begin{cases} \frac{\partial\eta}{\partial t} + \Delta\Psi = 0 \\ \frac{\partial\Psi}{\partial t} = 0. \end{cases}$$

En intégrant la première équation en temps, on trouve que  $-\Delta\Psi = \eta_1 - \eta_0$ . Les conditions aux bords sur  $\Psi$  sont les conditions de Neumann car on a posé que  $\partial\Psi/\partial\vec{n} = 0$ . Du coup la distance de Wasserstein vaut

$$\mathcal{W}_2(1 + \varepsilon\eta_0, 1 + \varepsilon\eta_1)^2 = \frac{1}{2} \iint_{[0,1] \times \Omega} \rho_\varepsilon |\nabla\Phi_\varepsilon|^2 = \frac{\varepsilon^2}{2} \iint_{[0,1] \times \Omega} |\nabla\Psi(x)|^2 + \mathcal{O}(\varepsilon^2) = \frac{\varepsilon^2}{2} \|\eta_1 - \eta_0\|_{-1} + \mathcal{O}(\varepsilon^2). \quad \blacksquare$$

**Le produit scalaire associé.** On définit le produit scalaire et la norme de deux vecteurs  $\eta, \eta'$  appartenant à  $T_\rho\mathcal{P}$  comme

$$\langle \eta, \eta' \rangle_{-1} = \int_\Omega \nabla\Delta^{-1}\eta \cdot \nabla\Delta^{-1}\eta', \quad \|\eta\|_{-1}^2 = \int_\Omega |\nabla\Delta^{-1}\eta|^2$$

avec toujours  $\Delta^{-1}$  dénotant l'inverse du Laplacien-Neumann.

### 2.8.4 Barycentre pour la distance $\mathcal{H}^{-1}$

Un simple calcul montre que le barycentre de  $\rho_0$  et  $\rho_1$  pour la distance  $\mathcal{H}^{-1}$  est

$$\rho = \frac{\rho_0 + \rho_1}{2}.$$

En effet, le minimiseur  $\rho$  de

$$\|\rho - \rho_0\|_{-1}^2 + \|\rho - \rho_1\|_{-1}^2 = \int_{\Omega} |\nabla \Delta^{-1}(\rho - \rho_0)|^2 + |\nabla \Delta^{-1}(\rho - \rho_1)|^2$$

vérifie la condition d'Euler-Lagrange  $\Delta^{-1}(2\rho - \rho_0 - \rho_1) = 0$ . C'est-à-dire que le barycentre est exactement le même que pour la distance euclidienne. Ainsi, même si cette distance mesure les mêmes choses que  $\mathcal{W}_2$ , le barycentre pour cette distance est complètement différent. Ce sera une limitation dans nos applications comme nous allons le voir dans la prochaine partie.

## 2.9 Conclusion

Pour conclure ce chapitre, nous faisons le rapprochement avec l'assimilation de données décrite dans le chapitre 1. Nous y avons dit que nous aimerions écrire une assimilation variationnelle de données de type images ou denses (avec une cohérence spatiale), lorsqu'il y a des erreurs de position. Ces erreurs de position seraient à l'origine de problèmes d'amplitude. Pour corriger ce problème, nous utiliserons la distance de Wasserstein capable d'interpoler des données localisées et mal placées, sans altérer l'amplitude des données. Nous utiliserons aussi la distance  $\mathcal{H}^{-1}$  qui ressemble en certains points à la distance de Wasserstein. Enfin, la distance  $WF$  sera utilisée pour le contrôle de variables positives dont la masse n'est pas imposée.

Numériquement, nous ferons des expériences en 1D pour lesquels la distance de Wasserstein est explicite. Les minimisations de la distance de Wasserstein et par extension de la distance  $WF$  seront réalisées en calculant leurs gradients dont nous avons fourni la valeur.





Deuxième partie

# Transport Optimal pour l'Assimilation de Données



## Chapitre 3

# Assimilation de données à l'aide du transport optimal

### Contenu du chapitre

3.1	État de l'art de l'assimilation de données et des problèmes inverses utilisant le transport optimal . . . . .	72
3.2	Présentation de l'assimilation de données utilisant le transport optimal . . . . .	73
3.2.1	Cas général . . . . .	73
3.2.2	Présentation des expériences qui vont suivre . . . . .	75
3.3	Minimisation d'une fonction coût de Wasserstein . . . . .	76
3.3.1	La méthode de descente de plus forte pente . . . . .	76
3.3.2	Descente du gradient euclidien de $\mathcal{J}(\rho_0)$ . . . . .	78
3.3.3	Descente du gradient de Wasserstein . . . . .	80
3.3.4	Formulation géodésique de la méthode de descente de plus forte pente . . . . .	83
3.3.5	Descente du gradient avec régularisation . . . . .	84
3.3.6	Résumé des méthodes de descente . . . . .	85
3.4	Descente de plus forte pente pour une fonction $\mathcal{H}^{-1}$ . . . . .	87
3.5	Les autres méthodes de descente du gradient . . . . .	88
3.5.1	La méthode du gradient conjugué . . . . .	88
3.5.2	Les méthodes de Newton et de quasi-Newton . . . . .	89
3.5.3	Implémentation numérique . . . . .	89

Nous avons conclu dans le chapitre d'assimilation de données que nous pouvions écrire de façon générale l'assimilation variationnelle de données comme la minimisation de la fonction coût :

$$\mathcal{J}(\mathbf{x}_0) := d_o(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + d_b(\mathbf{x}_0, \mathbf{x}_0^b)^2,$$

avec :

- des observations  $\mathbf{y}^{\text{obs}}$  de type images ou données denses ;
- une distance  $d_o$  capable de calculer la distance entre deux images, notamment en gérant les erreurs de position, donc une distance de Wasserstein ou une distance  $\mathcal{H}^{-1}$  présentées dans le chapitre précédent ;
- une distance  $d_b$ , de Wasserstein ou  $\mathcal{H}^{-1}$  s'il y a des erreurs de position dans l'ébauche. Dans ce cas là, il faudra que  $\mathbf{x}_0$  soit un élément de  $\mathcal{P}(\Omega)$  (ou un vecteur d'éléments de  $\mathcal{P}(\Omega)$ ). Nous discuterons de cette condition et de ce que l'on peut faire lorsque  $\mathbf{x}_0$  n'est pas dans  $\mathcal{P}(\Omega)$ .

Nous combinerons transport optimal et assimilation de données dans la section 3.2 et la minimisation numérique sera présentée dans la section 3.3. Avant, nous faisons une courte revue des quelques travaux scientifiques qui ont été déjà menées à ce sujet.

### 3.1 État de l'art de l'assimilation de données et des problèmes inverses utilisant le transport optimal

Assimiler les données en utilisant le transport optimal a rarement été fait. Nous citons cependant quelques initiatives.

**L'assimilation de données par réarrangements.** Dans [Douglas, 2002] l'auteur propose une méthode basée sur le transport optimal, et notamment les réarrangements, pour découpler les erreurs d'amplitude et de phase entre deux images.

Un réarrangement d'une image est une autre image où les pixels ont été changés de place. Ce réarrangement est défini *via* une permutation  $s : \{1, \dots, N_{\text{pix}}\} \rightarrow \{1, \dots, N_{\text{pix}}\}$  qui déplace le pixel  $i$  de l'image d'origine au pixel  $s(i)$  du réarrangement. À chaque réarrangement  $q$  d'une image  $q_1$  on associe un coût valant la distance de Wasserstein,  $\mathcal{W}_2(q, q_1)$  qui mesure le changement en position du réarrangement.

Soient deux images  $q_1$  et  $q_2$  dont on veut définir l'écart en amplitude et en position. On cherche le réarrangement  $q$  de  $q_1$  qui minimise :

$$\theta \mathcal{W}_2(q, q_1) + (1 - \theta) \|q_2 - q\|_2$$

La valeur du premier terme au minimum représente l'erreur de position, la valeur du deuxième représente l'erreur en amplitude.

En se basant sur cette idée de réarrangement, dans [Percival, 2008] est effectuée de l'assimilation de données en cherchant l'analyse parmi l'ensemble  $\mathcal{R}(\mathbf{x}_0^b)$  des réarrangements de l'ébauche, ce qui donne le problème suivant :

$$\min_{\mathbf{x}_0 \in \mathcal{R}(\mathbf{x}_0^b)} \|\mathbf{y}^{\text{obs}} - \mathcal{G}(\mathbf{x}_0)\|.$$

Cette méthode rappelle la méthode de réarrangement de la grille proposée dans [Ravela et al., 2007] et mentionnée dans la section 1.3.3, qui cherchait à déplacer l'ébauche.

Lorsqu'il y a effectivement des erreurs de position dans les données, cette méthode semble bien fonctionner pour du 3D-Var (lorsqu'il n'y a pas de modèle d'évolution impliqué). Pour du 4D-Var l'auteur précise que la méthode donne de bons résultats, même s'il rencontre des difficultés à minimiser efficacement sa fonction coût puisque le calcul du réarrangement est difficile.

**L'assimilation de données pour la sismologie par minimisation d'une fonction coût utilisant la distance de Wasserstein.** Comme on l'a expliqué plus tôt dans la Section 2.2.1, la distance de Wasserstein est convexe par rapport au déplacement et n'est pas sensible au bruit. Ces propriétés en font un bon élément pour traiter les données sismiques, comme proposé dans [Engquist et Froese, 2013]. Le modèle de propagation d'une onde  $u(t, x, z)$  est

$$\frac{\partial^2 u}{\partial t^2} - c(x, z)^2 \Delta u = 0 \tag{3.1.1}$$

avec  $c$  la conductivité du sol. Le but est de récupérer la valeur de  $c$ , notamment pour connaître la structure géologique souterraine. Pour ceci, des ondes sont créées à proximité par des explosions par exemple, et la réponse sismique est mesurée par des capteurs disposés à la surface  $z = 0$ . Classiquement, pour récupérer  $c$  on minimise :

$$\mathcal{J}(c) = \iint_{[0,T] \times \Omega} (u(c, t, x, z = 0) - u^{\text{obs}}(t, x))^2 dt dx \quad (3.1.2)$$

où  $u(c, t, x, z = 0)$  est la donnée de surface calculée à partir de  $c$  et du modèle (3.1.1). Il n'y a pas de terme d'ébauche puisque ce problème n'est pas vraiment sous-déterminé. En effet, une grille 2D  $(t, x)$  est utilisée pour retrouver la donnée 2D  $c(x, z)$ . La fonction  $\mathcal{J}(c)$  est non convexe car  $u$  est non linéaire en  $c$ , et possède même beaucoup de minima locaux sur lesquels la minimisation numérique s'arrêtera. Pour résoudre ce problème, les auteurs de [Engquist et al., 2016] proposent de retrouver  $c$  en minimisant :

$$\mathcal{J}(c) = \mathcal{W}_2(u_+^{\text{obs}}, u_+(c, z = 0)) + \mathcal{W}_2(u_-^{\text{obs}}, u_-(c, z = 0))$$

avec  $u_+$  et  $u_-$  les parties positives et négatives de  $u$  ( $u = u_+ - u_-$ ), normalisées pour que l'intégrale sur  $[0, T] \times \Omega$  vaille 1. Comme on l'a dit, cette fonction possède moins de minima locaux ce qui permet de récupérer le vrai  $c$  dans une expérience jumelle. L'autre travail important sur ce sujet est [Métivier et al., 2016]. Les auteurs proposent une distance heuristique qui permet de calculer directement la distance entre deux ondes, et minimisent ainsi

$$\mathcal{J}(c) = \widetilde{\mathcal{W}}_1(u^{\text{obs}}, u(c, z = 0)). \quad (3.1.3)$$

Ici  $\widetilde{\mathcal{W}}_1$  est définie en ajoutant une contrainte à la définition de  $\mathcal{W}_1$  pour pouvoir effectuer un transport optimal entre des fonctions non positives et de masses différentes. Cette distance est heuristique et est difficile à calculer numériquement, puisque c'est un problème de combinatoire, voir [Allain, 2016]. Cependant, l'analyse issue de la minimisation de la fonction coût (3.1.3) est de meilleure qualité que celle issue d'une fonction coût classique, voir la figure 3.1.

Les deux études montrées ici semblent bien fonctionner. La première basée sur le réarrangement est prometteuse mais est bloquée par le fait qu'il ait fallu passer par des réarrangements, difficiles à manipuler. La deuxième utilise une fonction coût basée sur une distance de Wasserstein et retrouve la conductivité du sol. Cependant, il n'y a pas de terme d'ébauche puisque le problème est bien posé. L'étude ne traite pas non plus les erreurs de position.

## 3.2 Présentation de l'assimilation de données utilisant le transport optimal

Nous présentons le contexte général de notre étude avant de présenter les résultats dans la suite du manuscrit.

### 3.2.1 Cas général

Soit un système physique, décrit sur un domaine spatial  $\Omega$  par l'état  $\mathbf{x}$ , un vecteur de différentes variables (un vecteur regroupant par exemple les variables de température, pression, vitesses, humidité...). On dispose d'un modèle  $\mathcal{M}$  décrivant l'évolution de  $\mathbf{x}$ . Les observations  $\mathbf{y}_i^{\text{obs}}$  sont des images, des captures de  $\rho(t, x)$ , une variable qui appartient ou non à l'état. Par exemple en océanographie  $\rho$  peut être un traceur (le phytoplancton) qui n'appartient pas à l'état, ou ce peut être la couverture nuageuse en météorologie qui appartient elle à l'état.

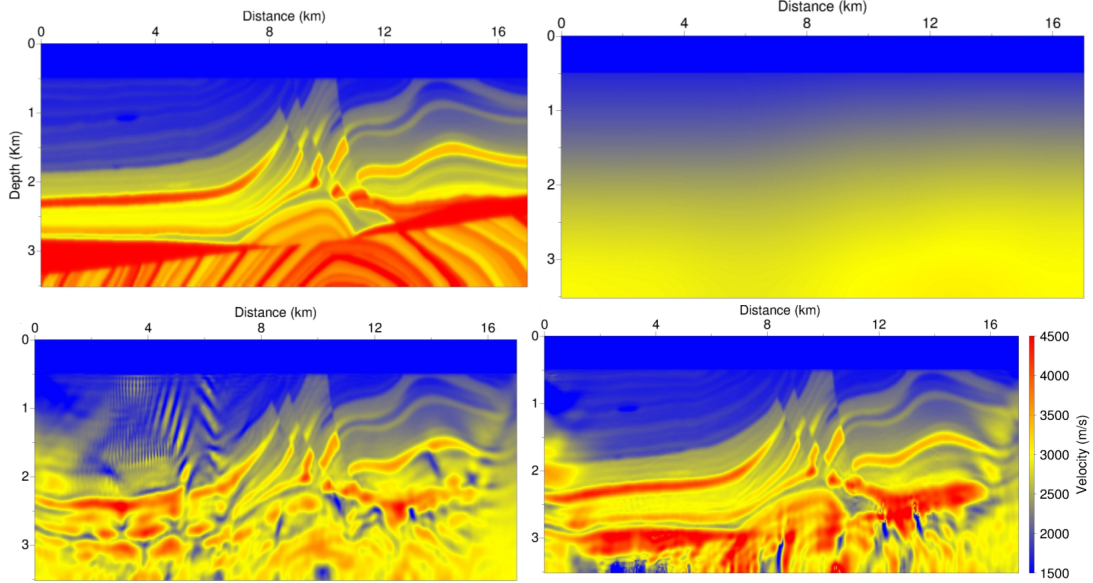


FIGURE 3.1 – Dans chaque image est représentée la vitesse  $c$  (en couleur) en fonction de la largeur  $x$  (en abscisse) et de la profondeur  $z$  (en ordonnée). En *haut à gauche* : solution exacte, en *haut à droite* : initialisation de la minimisation, en *bas à gauche* : solution issue de (3.1.2), en *bas à droite* : solution issue de (3.1.3). Source : [Métivier et al., 2016].

Les observations sont définies sur l'espace  $\Omega_o$  qui peut être différent de  $\Omega$  en pratique. De plus, les observations sont supposées être des éléments de  $\mathcal{P}(\Omega_o)$ , ce qu'on peut obtenir après normalisation. L'idée est d'assimiler les observations de type images à l'aide de la distance de Wasserstein : si nous écrivons la fonction coût comme

$$\mathcal{J}(\mathbf{x}_0) = \mathcal{J}^{\text{obs}}(\mathbf{x}_0) + \mathcal{J}^b(\mathbf{x}_0)$$

c'est-à-dire comme la somme de l'écart aux observations  $\mathcal{J}^o$  et de l'écart à l'ébauche  $\mathcal{J}^b$ , alors nous écrivons que

$$\mathcal{J}^{\text{obs}}(\mathbf{x}_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(\mathbf{x}_0), \mathbf{y}_i^{\text{obs}})^2.$$

Le but est à la fin de trouver une analyse  $\mathbf{x}_0$  qui soit "proche en position" des observations de type images. L'opérateur  $\mathcal{G}_i$  calcule l'équivalent de l'observation  $\mathbf{y}_i^{\text{obs}}$ .

Le terme d'écart à l'ébauche  $\mathcal{J}_b$  est aussi à définir. En toute généralité,  $\mathcal{J}_b$  vaut :

$$\mathcal{J}^b(\mathbf{x}_0) = d_b(\mathbf{x}_0, \mathbf{x}_0^b)^2 = \sum_j d_b^j(\mathbf{x}_0^j, \mathbf{x}_0^{j,b})^2$$

avec  $\mathbf{x}_0^j$  la variable numéro  $j$  de l'état  $\mathbf{x}_0$ , et  $d_b^j$  la distance utilisée pour cette variable. Notre volonté est de prendre des distances de Wasserstein, afin de traiter les erreurs de position. Nous ne pouvons utiliser cette distance pour toutes les variables, seulement pour celles appartenant à  $\mathcal{P}(\Omega)$ . Nous pouvons utiliser la distance  $WF$  pour les variables seulement positives comme la température. Pour les autres types de variable, soit il faudra les transformer, soit les laisser telles quelles et utiliser une distance  $\mathcal{L}^2$  par exemple. Nous étudierons ceci pour les variables de type vitesse qui ne sont pas des fonctions positives.

### 3.2.2 Présentation des expériences qui vont suivre

Nous allons effectuer quelques expériences pour tester une fonction coût  $\mathcal{J}$  écrite à l'aide de la distance de Wasserstein. Ces expériences seront purement numériques : le modèle numérique  $\mathcal{M}$  choisi sera supposé parfait (aucune différence entre la réalité et l'évolution décrite par ce modèle), ce qui est une hypothèse forte mais c'est une première étape d'expérimentation. Nous parlons alors d'expériences jumelles :

#### Définition 3.2.1 (*Expériences jumelles*)

Dans les expériences numériques qu'on appelle des *expériences jumelles*, le modèle numérique est supposé parfait et les observations utilisées seront créées à l'aide de celui-ci. Un état vrai  $\mathbf{x}_0^t$  est choisi, et les observations seront créées par intégration du modèle :  $\mathbf{y}^{\text{obs}} := \mathcal{G}(\mathbf{x}_0^t)$ . Les observations seront au besoin bruitées, par réalisme. Après avoir créé les observations, nous supposerons que nous ne connaissons pas  $\mathbf{x}_0^t$ . Ce type d'expérience est souvent utilisé pour tester une méthode nouvelle, avant de l'appliquer à des données réelles.

Nous effectuerons des expériences par étape :

1. Premièrement, nous n'aurons qu'une seule variable à contrôler,  $\rho_0$ , appartenant à  $\mathcal{P}(\Omega)$ , et dont les observations  $\rho_i^{\text{obs}}$  appartenant à  $\mathcal{P}(\Omega)$  seront directement issues :  $y_i^{\text{obs}} = \mathcal{M}_i(\rho_0)$ . La notation  $\mathcal{M}_i$  est utilisée pour définir l'intégration du modèle entre les temps  $t = 0$  et  $t = t_i$ . La fonction coût s'écrira alors

$$\mathcal{J}(\rho_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(\rho_0), \rho_i^{\text{obs}})^2 + \omega_\rho \mathcal{W}_2(\rho_0, \rho_0^b)^2$$

avec  $\omega_\rho$  un poids sur le terme d'ébauche. Le modèle  $\mathcal{M}$  sera une équation de continuité

$$\mathcal{M}: \frac{\partial \rho}{\partial t} + \text{div}(\rho u) = 0$$

avec  $u$  la vitesse, tangente aux bords de  $\Omega$ , et peut éventuellement dépendre de  $\rho$  comme dans le modèle de Saint-Venant que nous utiliserons.

Ce type d'expérience simple sera une première étape dans la compréhension du comportement de  $\mathcal{J}$ , notamment en cas d'erreurs de position. Ceci sera étudié dans le chapitre 4.

2. Dans une deuxième série d'expérience, nous aurons plusieurs variables à contrôler à l'aide d'observations  $\mathbf{y}_i^{\text{obs}}$  issues directement ou non de l'ensemble des variables de  $\mathbf{x}^t$ , mais observées sur tout le domaine  $\Omega$  (alors,  $\Omega_o = \Omega$ ). La fonction coût à écrire sera

$$\mathcal{J}(\mathbf{x}_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(\mathbf{x}_0), y_i^{\text{obs}})^2 + d_b(\mathbf{x}_0, \mathbf{x}_0^b)^2.$$

Cette série d'expérience a pour but :

- de contrôler plusieurs variables à l'aide de la distance de Wasserstein ;
- de montrer la capacité dans ce cas à gérer les erreurs de position ;
- de montrer la capacité à utiliser des variables indirectement observées ;
- d'étudier le choix du terme d'ébauche.

Ces expériences utiliseront un modèle de Saint-Venant et seront étudiées dans le chapitre 5.

3. Nous étudierons sommairement dans la section 6.2.4 le cas où le domaine entier n'est pas observé,  $\Omega^{\text{obs}} \neq \Omega$ , par exemple dans le cas d'occultations.



Les distances  $WF$  et  $\mathcal{H}^{-1}$  (définie à l'aide du laplacien-Neumann) seront brièvement étudiées.

Avant de présenter les expériences et les résultats, il faut savoir comment minimiser la fonction coût  $\mathcal{J}$ , notamment lorsque celle-ci utilise la distance de Wasserstein. C'est ce que nous étudions maintenant.

Enfin, notons que la prochaine section et la première série d'expérience a fait l'objet d'un article soumis [Feyeux et al., 2016].

### 3.3 Minimisation d'une fonction coût de Wasserstein

Nous nous intéressons dans cette partie à la minimisation d'une fonction coût écrite uniquement à l'aide de la distance de Wasserstein,

$$\mathcal{J}(\rho_0) = \mathcal{W}_2(\mathcal{M}(\rho_0), \rho^{\text{obs}})^2 + \omega_b \mathcal{W}_2(\rho_0, \rho_0^b)^2 \quad (3.3.1)$$

définie sur  $\mathcal{P}(\Omega)$ , où  $\mathcal{M}: \mathcal{P}(\Omega) \rightarrow \mathcal{P}(\Omega)$  est un modèle préservant l'appartenance à  $\mathcal{P}(\Omega)$ .

Il existe plusieurs types de minimisation d'une fonction générale, toutes itératives. On peut penser à la minimisation stochastique qui itère de façon plus ou moins aléatoire, ou à la minimisation sans gradient qui testera plusieurs directions de descente. Plutôt, nous avons choisi d'utiliser les méthodes de descente avec gradient qui sont classiquement utilisées en 4D-Var.

Nous allons donc employer des méthodes générales de descente avec gradient mais qui devront être adaptées à la minimisation avec la distance de Wasserstein, sur  $\mathcal{P}(\Omega)$ . En effet, avec ces méthodes classiques la convergence peut être longue et l'appartenance à  $\mathcal{P}(\Omega)$  n'est pas acquise. Ce problème sera contourné en utilisant une méthode de descente naturelle pour des fonctions définies avec la distance de Wasserstein.

Pour bien voir les particularités de la minimisation avec la distance de Wasserstein, nous allons étudier d'abord l'algorithme le plus simple : la descente de plus forte pente. Nous verrons ensuite dans la section 3.5 d'autres méthodes de descente plus complexes mais plus efficaces et plus utilisées, se basant sur l'algorithme de descente de plus forte pente.

#### 3.3.1 La méthode de descente de plus forte pente

La méthode de plus forte pente, comme les autres méthodes avec gradient, trouve le minimum en créant une suite  $(\rho_0^n)$ , de telle sorte que la suite  $(\mathcal{J}(\rho_0^n))$  soit décroissante et décroisse jusqu'au minimum, et que  $(\rho_0^n)$  converge vers le minimiseur. L'itération est de la forme

$$\rho_0^{n+1} = \rho_0^n + \alpha d^n$$

avec  $\alpha > 0$  un scalaire et  $d^n$  une direction de descente à trouver, qui dépendra du gradient de  $\mathcal{J}$ , nous étudierons ceci en détail. Le choix de  $d^n$  est le plus compliqué, mais une fois qu'il est trouvé, le choix de  $\alpha > 0$  sera facile puisque ce n'est qu'un scalaire :

#### Remarque 3.3.1 (Choix de $\alpha$ )

Une fois que  $d^n$  est fixé, le choix de  $\alpha$  est trouvé par des méthodes de *line search* :

- On peut choisir le  $\alpha$  minimisant  $f(\alpha) := \mathcal{J}(\rho_0^n + \alpha d^n)$ . Le trouver est assez simple puisque c'est une fonction d'une seule variable. On utilise généralement la méthode de Powell [Powell, 1964] qui combine plusieurs autres méthodes de recherche et qui trouve le minimum très rapidement (en peu d'itérations).
- On peut ne pas aller jusqu'au minimum de  $f(\alpha)$ . Le plus courant est de s'arrêter lorsque

les critères de Wolfe sont vérifiés. Ces critères sont très efficaces et imposent au  $\alpha$  choisi de ne pas être trop proche de l'origine, et que la courbure  $y$  soit suffisamment faible. Voir [Nocedal et Wright, 2006] pour plus de détails.

Il ne nous reste qu'à choisir  $d^n$ . Lorsque l'on veut minimiser sur un espace vectoriel  $E$  muni du produit scalaire  $\langle \cdot, \cdot \rangle$ , on peut utiliser l'algorithme de descente de plus forte pente qui prend l'opposé du gradient comme valeur de  $d^n$  :

**Algorithme 3.3.2 (Descente de plus forte pente)**

Soit  $f(x)$  une fonction définie sur un espace vectoriel  $E$ , le gradient de  $f$  en  $x$  est défini comme étant l'élément  $\text{grad } f(x)$  tel que

$$\forall d \in E, \quad f(x + \varepsilon d) = f(x) + \varepsilon \langle d, \text{grad } f(x) \rangle + \mathcal{O}(\varepsilon). \quad (3.3.2)$$

L'algorithme de descente de plus forte pente est l'itération

$$x^{n+1} = x^n - \alpha \text{grad } f(x^n)$$

avec  $\alpha > 0$  à trouver (voir la Remarque 3.3.1).

L'opposé du gradient est choisie parce que c'est la direction  $d$  qui minimise le plus la forme linéarisée (3.3.2) de  $f$  parmi les directions  $d$  de norme 1,

$$\arg \min_{d: \|d\|=1} \langle d, \text{grad } f(x) \rangle = -\frac{1}{\|\text{grad } f(x)\|} \text{grad } f(x).$$

Nous aurons à minimiser une fonction non pas sur un espace vectoriel complet, mais seulement sur le convexe  $\mathcal{P}(\Omega)$ . Et lorsqu'on minimise une fonction  $f$  sur un espace convexe  $C \subset E$ , on doit se confronter au problème que  $x^{n+1} = x^n - \alpha \text{grad } f(x^n)$  peut ne pas être dans  $C$ , quel que soit  $\alpha > 0$ . On ajoute alors une projection pour pouvoir rester dans  $C$ , ce qui donne l'algorithme du gradient projeté :

**Algorithme 3.3.3 (Méthode du gradient projeté)**

La méthode du gradient projeté est une méthode courante pour minimiser dans un espace convexe  $C$ . Pour résoudre

$$\inf_{x \in C} f(x),$$

la méthode du gradient projeté consiste à utiliser le gradient de  $f$ ,  $\text{grad } f$ , et à effectuer

$$x^{n+1} = P_C(x^n - \alpha \text{grad } f(x^n))$$

avec  $P_C$  la projection orthogonale sur  $C$ , définie par

$$P_C(x) = \arg \min_{\tilde{x} \in C} \|x - \tilde{x}\|.$$

Rappelons enfin que le gradient, la projection et donc les algorithmes présentés dépendent du produit scalaire choisi.

**Remarque 3.3.4**

Que  $E$  soit muni du produit scalaire  $\langle \cdot, \cdot \rangle_1$  ou de  $\langle \cdot, \cdot \rangle_2$ , la définition du gradient change. On aura

deux différents gradients  $\text{grad}_1 \mathcal{J}$  et  $\text{grad}_2 \mathcal{J}$  pour lesquels

$$\begin{aligned} \forall d \in E, \quad \mathcal{J}(x + \epsilon d) &= \mathcal{J}(x) + \epsilon \langle d, \text{grad}_1 \mathcal{J}(x) \rangle_1 + \mathcal{O}_{\epsilon \rightarrow 0}(\epsilon) \\ &= \mathcal{J}(x) + \epsilon \langle d, \text{grad}_2 \mathcal{J}(x) \rangle_2 + \mathcal{O}_{\epsilon \rightarrow 0}(\epsilon). \end{aligned}$$

On utilise d'abord le produit scalaire euclidien et l'algorithme du gradient projeté pour minimiser  $\mathcal{J}$  sur  $\mathcal{P}(\Omega)$ .

### 3.3.2 Descente du gradient euclidien de $\mathcal{J}(\rho_0)$

Nous allons utiliser la méthode du gradient projeté pour minimiser  $\mathcal{J}$  donnée en (3.3.1) qui vaut

$$\mathcal{J}(\rho_0) = \mathcal{W}_2(\mathcal{M}(\rho_0), \rho^{\text{obs}})^2 + \omega_b \mathcal{W}_2(\rho_0, \rho_0^b)^2.$$

Nous munissons  $\mathcal{P}(\Omega)$  du produit scalaire euclidien dans un premier temps. Étant donné que  $\mathcal{P}(\Omega)$  est convexe et à bord, l'algorithme du gradient projeté sera

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n - \alpha \text{grad} \mathcal{J}(\rho_0^n)).$$

Nous devons détailler deux termes ici : le gradient de  $\mathcal{J}$  et la projection sur  $\mathcal{P}(\Omega)$ .

#### **Théorème 3.3.5 (Gradient euclidien de $\mathcal{J}$ )**

On a pour tout  $\eta$  et  $\varepsilon > 0$  tels que  $\rho_0 + \varepsilon \eta \in \mathcal{P}(\Omega)$  :

$$\begin{aligned} \mathcal{J}(\rho_0 + \varepsilon \eta) &= \mathcal{J}(\rho_0) + \varepsilon \langle \eta, \Phi \rangle_2 + \mathcal{O}(\varepsilon) \\ &= \mathcal{J}(\rho_0) + \varepsilon \int_{\Omega} \eta(x) \Phi(x) \, dx + \mathcal{O}(\varepsilon). \end{aligned}$$

Le  $\Phi(x)$  est défini par

$$\Phi := \mathbf{M}^*[\rho_0] \cdot \Phi^{\text{obs}} + \omega_b \Phi^b,$$

où :

- $\Phi^{\text{obs}}$  est le potentiel de Kantorovich (qu'on a défini dans la définition 2.3.6 et dont les intrications avec les différents éléments du transport optimal sont résumées dans la Table 2.1) du transport entre  $\mathcal{M}(\rho_0)$  et  $\rho^{\text{obs}}$  ;
- $\Phi^b$  est le potentiel de Kantorovich du transport entre  $\rho_0$  et  $\rho_0^b$  ;
- et  $\mathbf{M}^*[\rho_0]$  est le modèle adjoint à  $\mathcal{M}$  en  $\rho_0$  (voir l'annexe B qui rappelle brièvement ce qu'est le modèle adjoint).

Ainsi, le gradient (pour le produit scalaire  $\mathcal{L}^2$ ) de  $\mathcal{J}$  est

$$\text{grad} \mathcal{J}(\rho_0) = \Phi$$

défini à une constante additive près.

**Démonstration :** Soit  $\eta$  tel que  $\rho_0 + \varepsilon \eta \in \mathcal{P}(\Omega)$  pour  $\varepsilon > 0$  suffisamment petit. Cela nécessite notamment que  $\int_{\Omega} \eta = 0$ . On calcule les variations infinitésimales. D'abord on constate qu'il faut différencier le modèle par rapport à  $\varepsilon$ ,

$$\mathcal{M}(\rho_0 + \varepsilon \eta) = \mathcal{M}(\rho_0) + \varepsilon \mathbf{M}[\rho_0] \cdot \eta + \mathcal{O}(\varepsilon).$$

avec  $\mathbf{M}[\rho_0]$  le modèle tangent en  $\rho_0$ . On doit aussi différencier la distance de Wasserstein ; la différentiation a été présentée dans le théorème 2.5.1 et fait apparaître le potentiel de Kantorovich. Ainsi, on a :

$$\begin{aligned} \mathcal{J}(\rho_0 + \varepsilon\eta) &= \frac{1}{2} \mathcal{W}_2(\mathcal{M}(\rho_0 + \varepsilon\eta), \rho^{\text{obs}})^2 + \frac{\omega_b}{2} \mathcal{W}_2(\rho_0 + \varepsilon\eta, \rho_0^b)^2 \\ &= \frac{1}{2} \mathcal{W}_2(\mathcal{M}(\rho_0) + \varepsilon \mathbf{M}[\rho_0].\eta + \mathcal{O}(\varepsilon), \rho^{\text{obs}})^2 + \frac{\omega_b}{2} \mathcal{W}_2(\rho_0 + \varepsilon\eta, \rho_0^b)^2 \\ &= \frac{1}{2} \mathcal{W}_2(\mathcal{M}(\rho_0), \rho^{\text{obs}})^2 + \varepsilon \langle \Phi^{\text{obs}}, \mathbf{M}[\rho_0].\eta \rangle_2 + \frac{\omega_b}{2} \mathcal{W}_2(\rho_0, \rho_0^b)^2 + \omega_b \varepsilon \langle \Phi^b, \eta \rangle_2 + \mathcal{O}(\varepsilon) \end{aligned}$$

avec  $\Phi^{\text{obs}}$  le potentiel de Kantorovich du transport entre  $\mathcal{M}(\rho_0)$  et  $\rho^{\text{obs}}$ , et  $\Phi^b$  le potentiel de Kantorovich du transport entre  $\rho_0$  et  $\rho_0^b$ . Aucun des deux ne dépend de  $\eta$ . En utilisant le modèle adjoint, on isole  $\eta$  grâce à la relation

$$\langle \Phi^{\text{obs}}, \mathbf{M}[\rho_0].\eta \rangle_2 = \langle \mathbf{M}^*[\rho_0].\Phi^{\text{obs}}, \eta \rangle_2$$

ce qui simplifie le calcul précédent en

$$\mathcal{J}(\rho_0 + \varepsilon\eta) = \mathcal{J}(\rho_0) + \varepsilon \langle \mathbf{M}^*[\rho_0].\Phi^{\text{obs}} + \omega_b \Phi^b, \eta \rangle_2 + \mathcal{O}(\varepsilon).$$

On pose  $\Phi = \mathbf{M}^*[\rho_0].\Phi^{\text{obs}} + \omega_b \Phi^b$ . C'est le gradient de  $\mathcal{J}$  mais défini à une constante près seulement, puisqu'on a aussi

$$\mathcal{J}(\rho_0 + \varepsilon\eta) = \mathcal{J}(\rho_0) + \varepsilon \langle \Phi + c, \eta \rangle_2 + \mathcal{O}(\varepsilon)$$

pour n'importe quelle constante  $c$ , car  $\int_{\Omega} \eta = 0$ . ■

Maintenant que nous avons le gradient, nous pouvons passer au calcul la projection sur  $\mathcal{P}(\Omega)$  :

### Proposition 3.3.6 (Projection sur $\mathcal{P}(\Omega)$ )

Soit  $f(x)$  une fonction de  $\Omega$ , n'appartenant pas forcément à  $\mathcal{P}(\Omega)$ . La projection orthogonale (au sens euclidien) de  $f$  sur  $\mathcal{P}(\Omega)$  est

$$P_{\mathcal{P}(\Omega)}(f) = (f - c)\chi_{[f-c \geq 0]}$$

avec  $\chi_{[f-c \geq 0]}$  la fonction caractéristique valant 1 où  $f - c \geq 0$  et 0 ailleurs, et  $c$  est la constante telle que l'intégrale de  $P_{\mathcal{P}(\Omega)}(f)$  sur  $\Omega$  vaille 1.

La valeur de  $c$  est trouvée numériquement très rapidement par dichotomie, étant donné que pour  $f$  continue,  $c \mapsto \int_{\Omega} (f - c)\chi_{[f-c \geq 0]} dx$  est une fonction croissante continue de  $c$ .

Nous démontrons la proposition 3.3.6 :

**Démonstration :** Soit  $f$  hors de  $\mathcal{P}(\Omega)$  et  $P$  sa projection sur  $\mathcal{P}(\Omega)$ , qui existe et est unique. La projection vérifie

$$\frac{1}{2} \|f - P\|^2 = \min_{g \in \mathcal{P}(\Omega)} \frac{1}{2} \|f - g\|^2 = \min_{g \geq 0, \int_{\Omega} g dx = 1} \frac{1}{2} \|f - g\|^2.$$

La contrainte  $\int_{\Omega} g dx = 1$  dans le minimum peut être passée par un multiplicateur de Lagrange,  $\lambda$ , qui est une constante,

$$\inf_{g \geq 0, \int_{\Omega} g dx = 1} \frac{1}{2} \|f - g\|^2 = \inf_{g \geq 0} \sup_{\lambda \in \mathbb{R}} \underbrace{\frac{1}{2} \|f - g\|^2 + \lambda \left( \int_{\Omega} [g - 1] dx \right)}_{\mathcal{L}(g, \lambda)}.$$

On a donc un lagrangien  $\mathcal{L}$  dont il faut trouver le point-selle. Par les variations infinitésimales, comme  $P$  est au point-selle, et en notant  $c$  la valeur de  $\lambda$  au point-selle, on a

$$\forall p \text{ tel que } (\forall \epsilon \in \mathbb{R}_+ \text{ assez petit, } P + \epsilon p \geq 0), \quad \mathcal{L}(P + \epsilon p, c) \geq \mathcal{L}(P, c)$$

qui implique, en prenant la limite  $\epsilon \rightarrow 0^+$ ,

$$\forall p \text{ positif là où } P = 0, \quad \int_{\Omega} p [P - f + c] \geq 0$$

ce qui caractérise  $P$ , puisqu'on a pour presque tout  $x$  de  $\Omega$ ,

$$\begin{cases} P(x) = 0 \Rightarrow f(x) - c \leq 0 \\ P(x) > 0 \Rightarrow P(x) = f(x) - c. \end{cases}$$

On montre alors facilement que  $P$  est bien égal à  $(f - c)\chi_{[f-c \geq 0]}$ . ■

### Algorithme 3.3.7 (*Gradient euclidien projeté*)

L'itération de la méthode du gradient projeté est donc

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n - \alpha \Phi^n)$$

avec  $\Phi^n = \mathbf{M}^*[\rho_0^n].\Phi^o + \omega_b \Phi^b$ , les potentiels  $\Phi^b$  et  $\Phi^o$  sont explicités dans le Théorème 3.3.5 et dépendent de  $\rho_0^n$ . Nous avons dit que  $\Phi^n$  était défini à une constante près, mais le choix de la constante n'a pas d'influence grâce à la projection sur  $\mathcal{P}(\Omega)$ .

Nous pouvons remarquer que cette écriture présente une incohérence. En effet, si nous regardons le type des variables, le gradient euclidien  $\Phi$  est la somme de  $\Phi^b$  et de  $\mathbf{M}^*.\Phi^o$ , c'est-à-dire que  $\Phi$  est homogène à un potentiel de Kantorovich. On l'ajoute à  $\rho_0^n$ , qui est une densité et qui est donc d'un type de variable complètement différent. Le potentiel de Kantorovich est notamment beaucoup plus lisse, comme nous pouvons le voir sur la figure 3.2 où nous avons tracé les deux types. En ajoutant ainsi un potentiel à une densité, on ne pourra pas retrouver les détails du minimiseur.

Pour résoudre ce problème, nous utilisons à la place du produit scalaire euclidien celui de Wasserstein. Ceci est étudié dans la prochaine section.

### 3.3.3 Descente du gradient de Wasserstein

Nous allons effectuer la même chose mais avec le produit scalaire de Wasserstein, c'est-à-dire calculer le gradient de Wasserstein de  $\mathcal{J}$  et ensuite expliciter la descente liée. Nous avons défini le produit scalaire de Wasserstein (cf. la définition 2.4.1) pour  $\eta = -\text{div}(\rho_0 \nabla \Phi)$  et  $\eta' = -\text{div}(\rho_0 \nabla \Phi')$  deux éléments de  $T_{\rho_0} \mathcal{P}$ , par

$$\langle \eta, \eta' \rangle_W = \int_{\Omega} \rho_0 \nabla \Phi \cdot \nabla \Phi' dx.$$

Le gradient de  $\mathcal{J}$  est alors :

### Théorème 3.3.8 (*Gradient de Wasserstein*)

Soit  $\Phi$  le gradient euclidien de  $\mathcal{J}(\rho_0)$  valant (cf. le théorème 3.3.5) :

$$\Phi := \mathbf{M}^*[\rho_0].\Phi^{\text{obs}} + \omega_b \Phi^b.$$

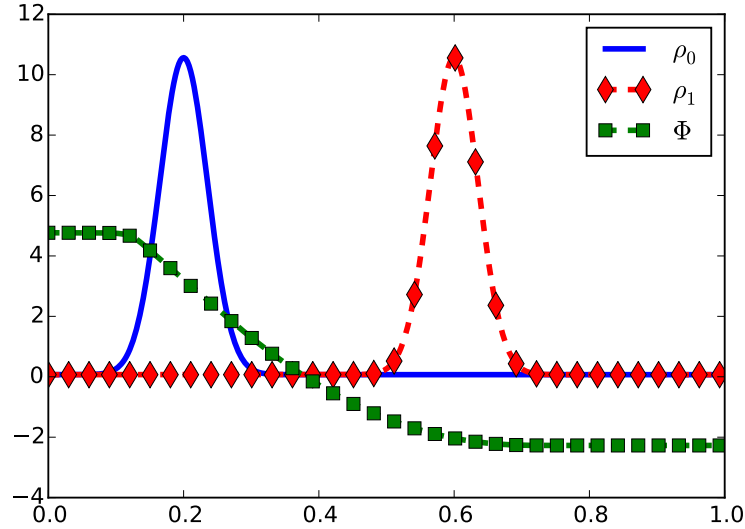


FIGURE 3.2 – Pour  $\rho_0$  et  $\rho_1$  on trace le potentiel de Kantorovich  $\Phi$  (normalisé) du transport du premier sur le second. Ces données ( $\rho$  et  $\Phi$ ) ne sont pas du même type.

En supposant que  $\rho_0 \frac{\partial \Phi}{\partial \vec{n}} = 0$  sur  $\partial\Omega$ , on a pour  $\eta \in T_{\rho_0} \mathcal{P}$  :

$$\mathcal{J}(\rho_0 + \epsilon\eta) = \mathcal{J}(\rho_0) + \epsilon \langle -\operatorname{div}(\rho_0 \nabla \Phi), \eta \rangle_W + \mathcal{O}(\epsilon)$$

avec le  $\Phi$  comme dans le Théorème 3.3.5 :

$$\Phi := \mathbf{M}^*[\rho_0] \cdot \Phi^{\text{obs}} + \omega_b \Phi^b.$$

Le gradient pour le produit scalaire de Wasserstein est alors

$$\operatorname{grad}_W \mathcal{J}(\rho_0) = -\operatorname{div}(\rho_0 \nabla \Phi).$$

qui appartient notamment à  $T_{\rho_0} \mathcal{P}$ .

**Démonstration :** Soit  $\eta \in T_{\rho_0} \mathcal{P}$  qui s'écrit à l'aide d'un potentiel  $\Psi$ ,

$$\eta = -\operatorname{div}(\rho_0 \nabla \Psi) \quad \text{avec} \quad \rho_0 \frac{\partial \Psi}{\partial \vec{n}} = 0 \quad \text{sur} \quad \partial\Omega.$$

On sait déjà que  $\mathcal{J}(\rho_0 + \epsilon\eta) = \mathcal{J}(\rho_0) + \epsilon \langle \Phi, \eta \rangle_2 + \mathcal{O}(\epsilon)$ . Le produit scalaire se réécrit

$$\begin{aligned} \langle \Phi, \eta \rangle_2 &= \int_{\Omega} \Phi \eta \, dx \\ &= \int_{\Omega} -\operatorname{div}(\rho_0 \nabla \Psi) \Phi \, dx \\ &= \int_{\Omega} \rho_0 \nabla \Psi \cdot \nabla \Phi \, dx - \int_{\partial\Omega} \underbrace{\Phi \rho_0 \frac{\partial \Psi}{\partial \vec{n}}}_{=0} \, dS \\ &= \langle \mu, \eta \rangle_W \quad \text{avec} \quad \mu := -\operatorname{div}(\rho_0 \nabla \Phi). \end{aligned} \tag{3.3.3}$$

La dernière égalité tient parce que  $\Phi$  vérifie aussi les conditions au bord de Neumann. Enfin, on peut remarquer que ce  $\mu$  appartient bien à  $T_{\rho_0} \mathcal{P}$ .


**Remarque 3.3.9 (Conditions aux bords de Neumann homogènes)**

Dans le théorème, nous avons supposé que

$$\rho_0 \frac{\partial \Phi}{\partial \bar{n}} = 0 \text{ sur } \partial \Omega \quad (3.3.4)$$

où  $\Phi = \mathbf{M}^*[\rho_0] \cdot \Phi^{\text{obs}} + \omega_b \Phi^b$ . Nous dirons que  $\Phi$  a une condition aux bords de Neumann homogène. Si cette condition n'est pas respectée,  $\mathcal{J}$  n'a pas de gradient puisque  $-\text{div}(\rho_0 \nabla \Phi)$  n'appartient plus à  $T_{\rho_0} \mathcal{P}$ .

Et cette condition n'est pas toujours respectée. En effet, même si les potentiels de Kantorovich  $\Phi^b$  et  $\Phi^{\text{obs}}$  ont des conditions aux bords de Neumann homogènes (cf. le corollaire 2.3.8), il n'y a pas de raison que dans le cas général  $\mathbf{M}^* \cdot \Phi^{\text{obs}}$  voie cette propriété conservée. Ainsi, l'hypothèse (3.3.4) n'est en général pas respectée. Nous utiliserons des exemples satisfaisant cette propriété.

Nous étudierons dans l'annexe C le cas où on n'a pas l'égalité (3.3.4) et verrons qu'on peut le résoudre en utilisant un produit scalaire qui prend en compte les valeurs aux bords. D'ici là, nous supposons toujours que le modèle adjoint préserve la condition aux bords de Neumann.

Maintenant que nous avons le gradient de Wasserstein, nous pouvons écrire la descente de plus forte pente de Wasserstein :

**Algorithme 3.3.10 (Descente de plus forte pente de Wasserstein)**

Avec ce gradient de Wasserstein, qui est :

$$\text{grad}_W \mathcal{J}(\rho_0^n) = -\text{div}(\rho_0^n \nabla \Phi^n) \quad \text{avec } \Phi^n = \mathbf{M}^*[\rho_0^n] \cdot \Phi^{\text{obs}} + \omega_b \Phi^b.$$

la descente devient :

$$\rho_0^{n+1} = \rho_0^n + \alpha^n \text{div}(\rho_0^n \nabla \Phi^n), \quad (3.3.5)$$

ou au besoin, en projetant dans  $\mathcal{P}(\Omega)$  :

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n + \alpha^n \text{div}(\rho_0^n \nabla \Phi^n)).$$

Cette version semble plus adaptée puisque puisque cette fois-ci  $\text{div}(\rho_0^n \nabla \Phi^n)$  est homogène à  $\rho_0^n$ . Cela peut se voir avec l'équation des géodésiques. En effet, l'équation de la géodésique de Wasserstein entre deux points  $\rho_0$  et  $\rho_1$  est, avec  $\Phi(x)$  le potentiel de Kantorovich entre  $\rho_0$  et  $\rho_1$  :

$$\begin{cases} \partial_t \rho + \text{div}(\rho \nabla \phi) = 0 \\ \partial_t \phi + \frac{|\nabla \phi|^2}{2} = 0 \\ \phi(0, x) = -\Phi(x). \end{cases} \quad (3.3.6)$$

Ici,  $\phi$  est donc une fonction de la même nature qu'un potentiel de Kantorovich. En discrétisant cette équation on a

$$\rho_1 \approx \rho_0 + \text{div}(\rho_0 \nabla \Phi)$$

ce qui montre bien que (3.3.5) fait intervenir des termes de même nature. Cependant, on peut recenser quelques problèmes :

- La nécessité de projeter sur  $\mathcal{P}(\Omega)$ , malgré l'appartenance du gradient dans  $T_{\rho_0^n} \mathcal{P}$ . Sans cette projection, la positivité ne serait pas assurée pour  $\alpha > 0$ . De plus, la projection ici est la

projection orthogonale au sens du produit scalaire  $\mathcal{L}^2$ , mais pas au sens de Wasserstein, qui elle n'est pas définie.

- Un autre problème est que  $\rho_0^{n+1}$  et  $\rho_0^n$  ont le même support, puisque  $\text{div}(\rho_0^n \nabla \Phi^n)$  a le même support que  $\rho_0^n$ . L'itération ne convergerait ainsi pas vers la solution puisque le support ne bougerait pas. Et dans le cas où  $\rho_0^n$  est relevé d'une petite constante, la convergence est en fait très lente.

Pour résoudre ces problèmes de convergence, on peut s'aider de la formulation des géodésiques : en effet, nous avons dit que (3.3.5) était la discrétisation en temps de l'équation des géodésiques (3.3.6). Cela nous inspire à écrire l'algorithme de descente de Wasserstein comme si la descente suivait une géodésique.

### 3.3.4 Formulation géodésique de la méthode de descente de plus forte pente

Au lieu d'écrire la formulation linéaire (en  $\Phi^n$ ) :

$$\rho_0^{n+1} = \rho_0^n + \alpha^n \text{div}(\rho_0^n \nabla \Phi^n),$$

nous allons écrire la formulation géodésique (non-linéaire), équivalente à la formulation linéaire pour  $\alpha \rightarrow 0$  :

#### **Algorithme 3.3.11** (*Descente du gradient de Wasserstein - forme transport*)

Soit  $\Phi^n$  le gradient euclidien de la fonction  $\mathcal{J}(\rho_0^n)$ . La formulation géodésique de la descente dans la direction du gradient de Wasserstein est :

$$\rho_0^{n+1} = (I - \alpha^n \nabla \Phi^n) \# \rho_0^n. \quad (3.3.7)$$

qui sont décrites par des équations semblables au système (3.3.6) (cf. le corollaire (2.3.10) page 53).

Étant donné que cette méthode est équivalente à la précédente pour  $\alpha^n = 0$ , on s'assure que  $\mathcal{J}(\rho_0^{n+1}) < \mathcal{J}(\rho_0^n)$  pour  $\alpha > 0$  suffisamment petit. Cette version présente de plus quelques avantages par rapport à la version précédente :

- Nous avons convergence en une itération si le modèle utilisé est simple. Par exemple pour  $\mathcal{M} = Id$ , la fonction coût à minimiser est

$$\mathcal{J}(\rho_0) = \mathcal{W}_2(\rho_0, \rho^{\text{obs}})^2 + \mathcal{W}_2(\rho_0, \rho_0^b)^2$$

et le minimum est au centre de la géodésique reliant  $\rho_0^b$  et  $\rho^{\text{obs}}$ . Ainsi, en partant de n'importe quel  $\rho_0^0$ , le gradient euclidien est

$$\Phi := \Phi^{\text{obs}} + \Phi^b$$

avec  $\Phi^{\text{obs}}$  le potentiel de Kantorovich entre  $\rho_0^0$  et  $\rho^{\text{obs}}$  et  $\Phi^b$  celui entre  $\rho_0^0$  et  $\rho_0^b$ . L'itération (3.3.7) converge en une seule itération, avec  $\alpha = 1/2$ . C'est ce qui est utilisé dans [Bigot et Klein, 2015] pour trouver le barycentre d'un nombre infini de densités.

- Dans la figure 3.3 nous comparons les deux itérations : la formulation linéaire (3.3.5) et la formulation géodésique (3.3.7), et nous pouvons en tirer diverses conclusions pratiques. D'abord les supports de  $\rho_0^{n+1}$  et de  $\rho_0^n$  sont différents, contrairement à la version précédente. De plus, on assure que  $\rho_0^{n+1}$  est dans  $\mathcal{P}(\Omega)$  sans avoir à faire de projection.



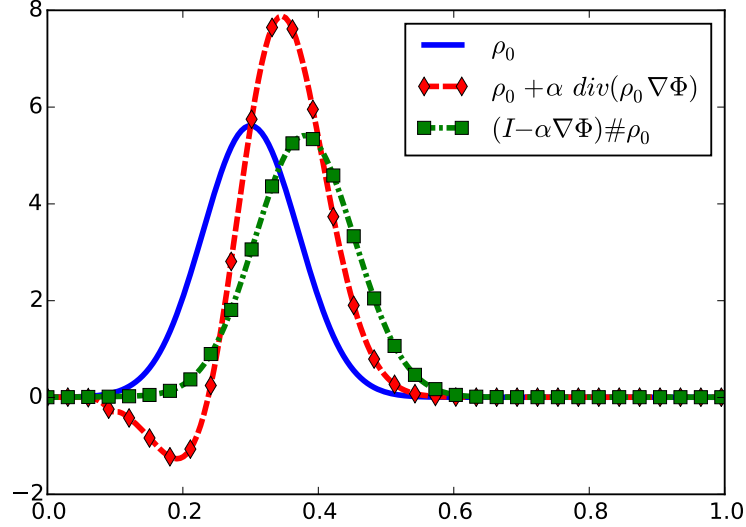


FIGURE 3.3 – Différence entre la formulation linéaire (3.3.5) qui donne la courbe rouge, et la formulation géodésique (3.3.7) qui donne la courbe verte.

Nous rencontrons tout de même des problèmes avec cette itération. Même si  $\rho_0^{n+1}$  est toujours dans  $\mathcal{P}(\Omega)$ , il se peut que l'itération (3.3.7) ne soit pas défini pour tout  $\alpha > 0$ . Nous avons parfois remarqué dans l'implémentation numérique que cette méthode fait face à des instabilités, l'apparition de Diracs notamment. Nous expliquons et corrigeons ce problème dans la section suivante.

### 3.3.5 Descente du gradient avec régularisation

Pour comprendre pourquoi cette dernière itération (3.3.7) n'est pas définie pour tout  $\alpha$ , nous pouvons écrire les équations continues calculant  $\rho(t) := (I - t\nabla\Phi)\#\rho_0$ ,

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho \nabla \phi) = 0, & \rho(0) = \rho_0 \\ \partial_t \phi + \frac{|\nabla \phi|^2}{2} = 0, & \phi(0) = -\Phi. \end{cases}$$

La deuxième équation est l'équation de Burgers potentielle. En posant  $v = \nabla \phi$ ,  $v$  résout l'équation de Burgers  $\partial_t v + v \cdot \nabla v = 0$ . Il est bien connu que cette équation de Burgers n'a pas de solution forte (continue) pour tout temps. En effet, des chocs peuvent arriver dès que les caractéristiques  $X_t(x)$  valant

$$X_t(x) = x - t\nabla\Phi(x)$$

ne sont plus inversibles en espace. Concrètement, cela crée dans  $\rho(t)$  des vagues qui se rencontrent et qui divergent, formant des Diracs ou des discontinuités. Une solution pratique est d'ajouter un peu de viscosité à l'équation de Burgers

$$\partial_t v + v \cdot \nabla v = \nu \Delta v$$

donnant l'équation de Burgers visqueuse dont les solutions fortes sont définies pour tout temps. En effet, par un simple changement de variable inversible (transformée de Cole-Hopf) cette équation se transforme en une équation de la chaleur.

La version potentielle (avec  $\phi$ ) de l'itération avec viscosité s'écrit alors :

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho_0 \nabla \phi) = 0, & \rho(0) = \rho_0 \\ \partial_t \phi + \frac{|\nabla \phi|^2}{2} = \nu \Delta \phi, & \phi(0) = -\Phi. \end{cases}$$

L'algorithme avec ce terme de diffusion s'écrit :

**Algorithme 3.3.12 (Formulation géodésique avec régularisation)**

$\rho_0^{n+1}$  est la solution à  $t = \alpha^n$  de

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho \nabla \phi) = 0, & \rho(0) = \rho_0^n \\ \partial_t \phi + \frac{|\nabla \phi|^2}{2} = \nu \Delta \phi, & \phi(0) = -\Phi^n, \end{cases} \quad (3.3.8)$$

itération que nous noterons :

$$\rho_0^{n+1} = (I - \alpha \nabla \Phi^n) \#^\nu \rho_0^n.$$

Cette méthode présente de nombreux avantages :

- D'abord, comme on l'a dit, cette itération est définie pour tout temps. En plus de l'assurance que  $\rho_0^{n+1}$  est dans  $\mathcal{P}(\Omega)$ , il n'y a pas création de pics qui auraient aussi pour conséquence de coincer la minimisation dans un minimum local.
- Cette itération est toujours équivalente à la descente de Wasserstein linéaire (3.3.5) dans le cas  $\alpha$  infinitésimal. En effet, avec une discrétisation Euler explicite de (3.3.8), on a toujours

$$\rho(\alpha) = \rho_0^n + \alpha \operatorname{div}(\rho_0^n \nabla \Phi^n).$$

Nous sommes alors assurés d'avoir  $\mathcal{J}(\rho_0^{n+1}) < \mathcal{J}(\rho_0^n)$  pour  $\alpha > 0$  petit.

Ce terme de diffusion apporte cependant quelques défauts :

- Nous perdons la convergence en un seul coup dans le cas de modèles simples. En effet, (3.3.8) n'est plus l'équation d'une géodésique.
- Cette itération ne fait pas apparaître les détails fins puisque ceux-ci sont lissés ; il ne faudra pas prendre  $\nu$  trop grand. Concrètement, nous avons choisi de le faire décroître au fur et à mesure des itérations de la descente, et de le mettre égal à 0 à partir d'un moment.

### 3.3.6 Résumé des méthodes de descente

On a défini quatre différentes méthodes de descente du gradient. D'abord il y a celle avec le gradient euclidien, ensuite celle avec le gradient de Wasserstein, puis la formulation géodésique et enfin celle avec la régularisation. Ces quatre méthodes s'écrivent :

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n - \alpha^n \Phi^n) \quad (\text{DG2})$$

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n + \alpha^n \operatorname{div}(\rho_0^n \nabla \Phi^n)) \quad (\text{DG})$$

$$\rho_0^{n+1} = (I - \alpha^n \nabla \Phi^n) \# \rho_0^n \quad (\text{DG}\#)$$

$$\rho_0^{n+1} = (I - \alpha^n \nabla \Phi^n) \#^\nu \rho_0^n \quad (\text{DG}\#\nu)$$

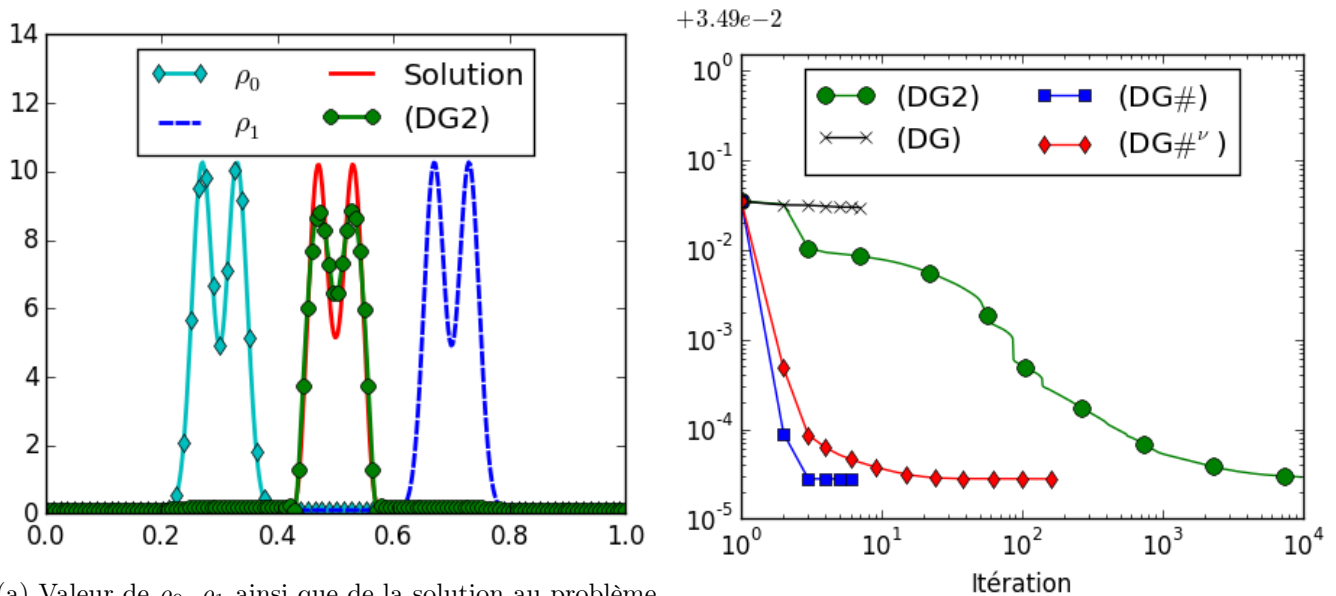
Nous illustrons la performance de ces algorithmes sur un exemple simple. La fonction coût à minimiser est

$$\mathcal{J}(\rho) = \mathcal{W}_2(\rho, \rho_0)^2 + \mathcal{W}_2(\rho, \rho_1)^2 \quad (3.3.10)$$

et est minimisée à l'aide des différents algorithmes. Le point de départ des algorithmes est toujours  $\rho_0$  et l'algorithme s'arrête lorsque le gradient de  $\mathcal{J}$  devient trop petit (de norme inférieure à  $10^{-10}$ ). Pour l'algorithme (DG $\#^\nu$ ), nous avons gardé un  $\nu$  constant.

La figure 3.4a affiche  $\rho_0$  et  $\rho_1$ , deux densités localisées, et le barycentre situé au milieu. Nous affichons sur la figure 3.4b la vitesse de convergence en traçant la valeur de  $\mathcal{J}$  en fonction des itérations. Les résultats des différents algorithmes sont :

- L'algorithme (DG2) n'a pas convergé au bout de 10 000 itérations, cf. l'analyse de la figure 3.4a où nous pouvons constater qu'elle ne coïncide pas exactement avec la solution. Le minimum est globalement retrouvé, mais la convergence n'est pas atteinte notamment parce qu'il manque les détails dans le gradient.
- L'algorithme (DG) donne une analyse présentant plein de pics (analyse non affichée). L'algorithme n'a pas convergé sûrement à cause du fait que la projection utilisée n'est pas une projection de Wasserstein (qui n'est elle pas définie).
- Les algorithmes (DG $\#$ ) et (DG $\#^\nu$ ) ont retrouvé la solution, et rapidement : respectivement en 6 et 100 itérations. L'algorithme (DG $\#^\nu$ ) converge un peu plus lentement à cause du terme de diffusion lissant le gradient et empêchant d'atteindre le minimum.



(a) Valeur de  $\rho_0$ ,  $\rho_1$  ainsi que de la solution au problème (3.3.10) (le barycentre de  $\rho_0$  et  $\rho_1$ ) et la solution donnée par (DG2) après 10 000 itérations.

(b) Valeur de  $\mathcal{J}$  en fonction du nombre d'itérations et de l'algorithme utilisé.

FIGURE 3.4 – Données de la minimisation de  $\mathcal{J}$  donné dans (3.3.10). À gauche (a) : les  $\rho_0$ ,  $\rho_1$ , le barycentre et aussi la solution donnée par (DG2) (la solution donnée par (DG $\#^\nu$ ) et (DG $\#$ ) sont égales au barycentre, et (DG) n'a pas convergé) ; à droite (b) : valeur de  $\mathcal{J}$  en fonction des algorithmes et de l'itération.

Concrètement dans la suite, pour minimiser  $\mathcal{J}$  écrite à l'aide de  $\mathcal{W}_2$ , nous utiliserons d'abord l'algorithme (DG $\#^\nu$ ) étant donné que comme on l'a dit, cette itération est définie pour tout  $\alpha > 0$ .

Nous faisons décroître le  $\nu$  et une fois que la solution est proche du minimum global, nous mettons  $\nu$  à 0 pour faire apparaître tous les détails de la solution.

### 3.4 Descente de plus forte pente pour une fonction $\mathcal{H}^{-1}$

Nous avons réalisé un travail de changement de gradient pour minimiser une fonction écrite à l'aide de la distance de Wasserstein. Nous allons à présent faire de même mais pour une fonction de  $\mathcal{P}(\Omega)$  écrite à l'aide de la distance  $\mathcal{H}^{-1}$ , en utilisant le gradient  $\mathcal{H}^{-1}$ . Ce travail sera beaucoup plus court car nous annonçons directement la méthode de descente de plus forte pente.

Nous rappelons que nous avons défini  $\mathcal{H}^{-1}$  dans la section 2.8 avec l'aide d'un laplacien-Neumann.

Soit la fonction coût  $\mathcal{H}^{-1}$  :

$$\mathcal{J}_{\mathcal{H}}(\rho_0) = \frac{1}{2} \|\mathcal{M}(\rho_0) - \rho^{\text{obs}}\|_{-1}^2 + \frac{\omega_b}{2} \|\rho_0 - \rho_0^b\|_{-1}^2.$$

#### Méthode 3.4.1 (Descente de plus forte pente pour $\mathcal{J}_{\mathcal{H}}$ )

Soit  $\eta$  une petite perturbation de  $\rho_0 \in \mathcal{P}(\Omega)$ . Soit  $d = \Delta^{-1}(\mathcal{M}(\rho_0) - \rho_0^b)$  l'écart avec les observations (avec condition aux bords de Neumann). Si

$$\frac{\partial \mathbf{M}^* . d}{\partial \vec{n}} = 0 \text{ sur } \partial \Omega \quad (3.4.1)$$

alors on a :

$$\mathcal{J}_{\mathcal{H}}(\rho_0 + \varepsilon \eta) = \mathcal{J}_{\mathcal{H}}(\rho_0) + \varepsilon \left\langle \Delta \mathbf{M}^* . d + \omega_b(\rho_0 - \rho_0^b), \eta \right\rangle_{-1} + \mathcal{O}(\varepsilon).$$

Nous trouvons donc le gradient  $\mathcal{H}^{-1}$  de  $\mathcal{J}_{\mathcal{H}}$  :

$$\text{grad}_{-1} \mathcal{J}_{\mathcal{H}} = \Delta \mathbf{M}^* . d + \omega_b(\rho_0 - \rho_0^b)$$

et la descente s'écrit

$$\rho_0^{n+1} = P_{\mathcal{P}(\Omega)}(\rho_0^n - \alpha^n \text{grad}_{-1} \mathcal{J}_{\mathcal{H}}(\rho_0^n)).$$

Nous trouvons notamment le même type de condition d'existence du gradient que dans le théorème 3.3.8 : la préservation de la condition d'être à condition aux bords de Neumann. Cette condition sera discutée dans l'annexe C.

**Démonstration du gradient :** Soit  $\eta$  une perturbation de  $\rho_0$ . Le terme d'écart à l'ébauche  $\mathcal{J}^b$  s'écrit

$$\begin{aligned} \mathcal{J}^b(\rho_0 + \varepsilon \eta) &= \frac{\omega_b}{2} \|\rho_0 + \varepsilon \eta - \rho_0^b\|_{-1}^2 \\ &= \mathcal{J}^b(\rho_0) + \varepsilon \omega_b \langle \eta, \rho_0 - \rho_0^b \rangle_{-1} + \mathcal{O}(\varepsilon). \end{aligned}$$

Le terme d'écart aux observations s'écrit lui :

$$\begin{aligned} \mathcal{J}^o(\rho_0 + \varepsilon \eta) &= \frac{1}{2} \|\mathcal{M}(\rho_0 + \varepsilon \eta) - \rho^{\text{obs}}\|_{-1}^2 \\ &= \mathcal{J}^o(\rho_0) + \varepsilon \underbrace{\langle \mathcal{M}(\rho_0) - \rho^{\text{obs}}, \mathbf{M} \eta \rangle_{-1}}_I + \mathcal{O}(\varepsilon) \end{aligned}$$

avec  $I$  valant, en prenant la notation  $d = \Delta^{-1}(\mathcal{M}(\rho_0) - \rho_0^b)$  :

$$\begin{aligned}
I &= \langle \mathcal{M}(\rho_0) - \rho^{\text{obs}}, \mathbf{M}\eta \rangle_{-1} = \int_{\Omega} \underbrace{\nabla \Delta^{-1}(\mathcal{M}(\rho_0) - \rho^{\text{obs}})}_d \cdot \nabla \Delta^{-1} \mathbf{M}\eta \\
&= - \int_{\Omega} d \mathbf{M}\eta \quad \text{parce qu'on a un laplacien-Neumann} \\
&= - \int_{\Omega} \mathbf{M}^* d \eta \\
&= - \int_{\Omega} (\Delta^{-1} \Delta \mathbf{M}^* d) (\nabla \cdot \nabla \Delta^{-1} \eta) \quad \text{en utilisant (3.4.1)} \\
&= \int_{\Omega} (\nabla \Delta^{-1} \Delta \mathbf{M}^* d) (\nabla \Delta^{-1} \eta) \quad \text{parce qu'on a un laplacien-Neumann} \\
&= \langle \Delta \mathbf{M}^* d, \eta \rangle_{-1}
\end{aligned}$$

Nous avons réussi à isoler le  $\eta$  dans  $\mathcal{J}^b$  et  $\mathcal{J}^o$ . La fin de la démonstration en découle directement. ■

### Remarque 3.4.2

Un problème de taille dans la formulation de la méthode est la projection  $\mathcal{H}^{-1}$  sur  $\mathcal{P}(\Omega)$ , définie pour  $f \notin \mathcal{P}(\Omega)$  par

$$P_{\mathcal{P}(\Omega)}(f) = \arg \inf_{\rho \in \mathcal{P}(\Omega)} \|f - \rho\|_{-1}.$$

Mais nous n'avons pas réussi à obtenir de formulation explicite de la projection. Sans formulation explicite, la valeur projection peut être obtenue par des méthodes numériques d'optimisation.

Nous avons utilisé la projection  $\mathcal{L}^2$  mais rien ne garantit la convergence de  $\rho_0^{n+1}$  vers le minimiseur de  $\mathcal{J}_{\mathcal{H}}$ ...

C'est pourquoi les fonctions coûts  $\mathcal{H}^{-1}$  seront délicates à utiliser sur  $\mathcal{P}(\Omega)$ .

## 3.5 Les autres méthodes de descente du gradient

La méthode de descente de plus forte pente est simple mais a le défaut de ne pas toujours être efficace. Outre le choix du produit scalaire qui n'est pas évident, il est à peu près clair que dans la plupart des cas, la direction de plus forte pente n'est pas la meilleure direction. D'autres méthodes ont été proposées, où les directions précédentes sont sauvegardées pour comprendre comment la fonction décroît, et affiner la direction actuelle et converger plus rapidement. Parmi ces méthodes il y a le gradient conjugué et des méthodes quasi-Newton se basant sur la méthode de Newton.

Nous avons vu qu'il fallait changer de produit scalaire dans la méthode de descente de plus forte pente pour effectuer efficacement la minimisation de  $\mathcal{J}$ . Nous présentons succinctement ces autres algorithmes et prétendons qu'il est possible de les utiliser en utilisant un produit scalaire de Wasserstein. Nous avons implémenté ces méthodes mais pour un produit scalaire  $\mathcal{H}^{-1}$ , plus simple d'utilisation.

### 3.5.1 La méthode du gradient conjugué

Le gradient conjugué se base sur la méthode de descente de plus forte pente, mais en gardant en mémoire les itérations précédentes pour comprendre la structure de la fonction, ce qui a pour effet d'accélérer la convergence. L'algorithme construit la direction de descente  $d^n$  à l'aide du gradient et de la direction précédente,

$$d^n = -\text{grad}\mathcal{J}(\rho_0^n) + \beta^n d^{n-1}$$

où  $\beta^n$  est une constante calculée, par exemple avec la formule de Fletcher-Reeves

$$\beta_{FR}^n = \frac{\|\text{grad}\mathcal{J}(\rho_0^n)\|^2}{\|\text{grad}\mathcal{J}(\rho_0^{n-1})\|^2}.$$

Il existe d'autres valeurs pour  $\beta^n$ , comme celle de Polak-Ribière,

$$\beta_{PR}^n = \frac{\langle \text{grad}\mathcal{J}(\rho_0^n), \text{grad}\mathcal{J}(\rho_0^n) - \text{grad}\mathcal{J}(\rho_0^{n-1}) \rangle}{\|\text{grad}\mathcal{J}(\rho_0^{n-1})\|^2}$$

ou celle qui réinitialise la valeur de  $\beta^n$  à chaque fois qu'elle devient négative,  $\beta^n = \max(0, \beta_{PR}^n)$ .

Dans cette méthode de gradient conjugué, le gradient et le produit scalaire utilisés ne sont pas nécessairement euclidiens. La théorie ne donne pas produit scalaire en particulier à utiliser et donc il est tout à fait possible d'utiliser des produits scalaire de Wasserstein.

### 3.5.2 Les méthodes de Newton et de quasi-Newton

Nous avons vu que la méthode de descente de plus forte pente était définie comme la minimisation de l'approximation d'ordre 1 de la fonction  $\mathcal{J}$ . Nous pouvons faire de même à l'ordre 2, approximation qui s'écrit :

$$\mathcal{J}(\rho_0 + \eta) = \mathcal{J}(\rho_0) + \underbrace{\langle \eta, \text{grad}\mathcal{J} \rangle + \frac{1}{2} \langle H(\rho_0) \cdot \eta, \eta \rangle}_{Q(\eta)} + \mathcal{O}(\|\eta\|^2)$$

avec  $H(\rho_0)$  la Hessienne, de  $\mathcal{J}$ . C'est une forme linéaire sur  $T_{\rho_0}\mathcal{P}$ . Le  $\eta$  minimisant la forme quadratique  $Q(\eta)$  résout

$$H(\rho_0) \cdot \eta = -\text{grad}\mathcal{J}(\rho_0).$$

C'est la méthode de Newton : l'incrément est donné par le minimiseur de la forme quadratique,

$$\rho_0^{n+1} = \rho_0^n - H(\rho_0)^{-1} \cdot \text{grad}\mathcal{J}(\rho_0).$$

C'est une méthode très performante et elle a la particularité de ne pas dépendre du produit scalaire choisi. En effet, en changeant de produit scalaire, on change le gradient et l'inverse de la Hessienne, mais pas le produit des deux. Il n'y a donc pas le problème de savoir quel gradient utiliser. Le seul inconvénient majeur est que l'on n'a généralement pas accès à la Hessienne et encore moins à son inverse.

C'est pourquoi les méthodes de quasi-Newton ont été écrites. Il est possible d'approcher la valeur de l'inverse de la Hessienne à l'aide de la valeur des gradients et de son évolution. Cet algorithme, nommé aussi BFGS (pour Broyden-Fletcher-Goldfard-Shanno), n'est pas décrit ici.

### 3.5.3 Implémentation numérique

Le stagiaire Benoît Kluger a travaillé sur ces méthodes de descente utilisant le produit scalaire  $\mathcal{H}^{-1}$ . Il a montré que pour minimiser des fonctions coûts utilisant la distance  $\mathcal{H}^{-1}$ , il valait souvent mieux utiliser le produit scalaire  $\mathcal{H}^{-1}$  plutôt que le produit scalaire  $\mathcal{L}^2$ . Ceci est montré dans la figure 3.5. Les résultats sont théoriquement les mêmes pour le produit scalaire  $\mathcal{W}_2$  à la place du produit scalaire  $\mathcal{L}^2$  pour des fonctions coûts écrites à l'aide de la distance de Wasserstein. Il faudra combiner la valeur de ces gradients avec des descentes de type (DG# $\nu$ ).

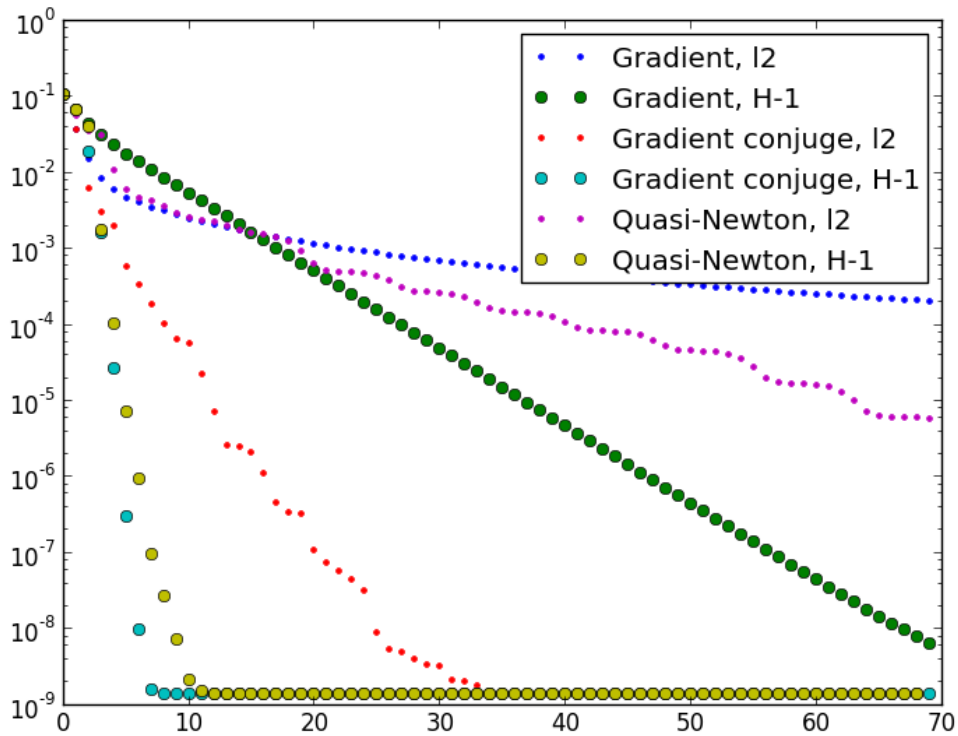


FIGURE 3.5 – Valeur d’une fonction coût écrite à l’aide de la distance  $\mathcal{H}^{-1}$  en fonction de l’itération et de la méthode utilisée, dépendant du produit scalaire  $\mathcal{L}^2$  (l2 dans la légende) et  $\mathcal{H}^{-1}$  (H-1 dans la légende). Ces méthodes sont la méthode de plus forte pente (gradient dans la légende), gradient conjugué et quasi-Newton.

# Chapitre 4

## Contrôle d'une densité de probabilité

### Contenu du chapitre

4.1	Choix du terme d'ébauche . . . . .	91
4.2	Description des expériences suivantes . . . . .	94
4.3	Contrôle d'une seule densité dans un modèle linéaire. . . . .	95
4.4	Assimilation dans un modèle de Saint-Venant . . . . .	95
4.5	Ajout d'un terme de régularisation . . . . .	97
4.6	Capacité à retrouver la forme de l'objet . . . . .	99
4.7	Observations soumises à du bruit . . . . .	102
4.8	Dans le cas de présence d'erreur d'amplitude seulement . . . . .	103
4.9	Distance Wasserstein-Fisher-Rao pour des densités positives . . . . .	105
4.10	Conclusion de ces contrôles d'une seule densité de probabilité . . . . .	106

Dans le chapitre 3, nous avons voulu écrire une fonction coût à l'aide de la distance de Wasserstein, notamment pour assimiler les images. Nous avons dit que nous allons procéder par étape. D'abord, et ce sera l'objet de ce chapitre, nous contrôlerons une unique variable qui est directement observée à sa sortie du modèle. Cette variable à assimiler sera  $\rho_0 \in \mathcal{P}(\Omega)$  et les observations seront notées  $\rho_i^{\text{obs}} \in \mathcal{P}(\Omega)$ .

Pour assimiler les données, nous allons utiliser la fonction coût :

$$\mathcal{J}(\rho_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{M}_i(\rho_0), \rho_i^{\text{obs}})^2 + \omega_b d_b(\rho_0, \rho_0^b).$$

Le choix la distance sur l'ébauche  $d_b$  sera discuté, notamment dans le cas où l'ébauche est sujette à des erreurs de position. Nous verrons aussi plus succinctement une distance  $\mathcal{H}^{-1}$  ou  $WF$  sur les observations.

Avant de présenter les expériences qui seront menées dans ce chapitre, nous étudions le choix de la distance  $d_b$  lorsqu'il y a des erreurs de position et lorsqu'il n'y a pas de modèle (3D-Var).

### 4.1 Choix du terme d'ébauche

Nous regardons dans cette première section un exemple de 3D-Var : il n'y a pas de modèle et l'observation  $\rho^{\text{obs}}$  est directement une capture de  $\rho_0^t$ . Les données (l'observation et l'ébauche)



seront sujettes à des erreurs de position, nous les prendrons décalées dans l'espace. L'assimilation variationnelle de données s'écrit dans le cas général comme la minimisation de

$$\mathcal{J}(\rho_0) = d_o(\rho_0, \rho^{\text{obs}})^2 + \omega_b d_b(\rho_0, \rho_0^b)^2$$

où il faut choisir les distances d'observation  $d_o$  et d'ébauche  $d_b$  et le poids  $\omega_b$ . Pour étudier l'influence du choix des distances, nous traçons sur la figure 4.1 les résultats de la minimisation pour plusieurs couples de distances, parmi les distances  $\mathcal{L}^2$ ,  $\mathcal{W}_2$  et  $\mathcal{H}^{-1}$ , et pour plusieurs valeurs de  $\omega_b$  allant de 0 (le minimiseur étant  $\rho_0^t$ ) à  $+\infty$  (le minimiseur étant  $\rho_0^b$ ). Le but est de retrouver  $\rho_0^t$  qui est placé entre l'ébauche et l'observation.

- On retrouve les barycentres classiques (euclidiens) dans les cas où  $(d_o, d_b) = (\mathcal{L}^2, \mathcal{L}^2)$  ou  $(d_o, d_b) = (\mathcal{H}^{-1}, \mathcal{H}^{-1})$  sur les figures 4.1a et 4.1b.
- De même pour  $(d_o, d_b) = (\mathcal{W}_2, \mathcal{W}_2)$  en 4.1c on retrouve les barycentres de type Wasserstein.
- Pour les couplages de distance faisant intervenir la distance euclidienne, par exemple pour  $(d_o, d_b) = (\mathcal{W}_2, \mathcal{L}^2)$  ou  $(d_o, d_b) = (\mathcal{H}^{-1}, \mathcal{L}^2)$  en 4.1d ou 4.1e, le minimiseur sera une diffusion de la partie mesurée par  $\mathcal{W}_2$  ou  $\mathcal{H}^{-1}$ , et une excroissance de celle mesurée par  $\mathcal{L}^2$ , sans correction de position, avec des défauts de largeur et des pertes de détails.

C'est une forme particulière, dont l'interprétation est donnée par les conditions d'Euler-Lagrange que remplit le minimiseur. Par exemple, dans le cas  $(d_o, d_b) = (\mathcal{H}^{-1}, \mathcal{L}^2)$ , la fonction coût à minimiser est :

$$\mathcal{J}(\rho_0) = \|\rho_0 - \rho^{\text{obs}}\|_{-1}^2 + \omega_b \|\rho_0 - \rho_0^b\|_2^2$$

et le gradient (euclidien) étant nul à l'analyse  $\rho_0^a$ , on a :

$$-\Delta^{-1}(\rho_0^a - \rho^{\text{obs}}) + \omega_b(\rho_0^a - \rho_0^b) = 0$$

et donc  $\rho_0^a$  résout le problème suivant :

$$(I - \omega_b \Delta) \rho_0^a = \rho^{\text{obs}} - \omega_b \Delta \rho_0^b.$$

Nous voyons dans le second membre que  $\rho^{\text{obs}}$  et  $\rho_0^b$  ne sont pas traités de la même manière :  $\rho^{\text{obs}}$  est mis au même niveau que le laplacien de  $\rho_0^b$ . Les densités ne sont pas comparées à des densités, cela est lié à pourquoi l'analyse a ces formes-ci. Ainsi, prendre  $\mathcal{L}^2$  pour l'ébauche lorsque l'on a choisi  $\mathcal{W}_2$  ou  $\mathcal{H}^{-1}$  pour les observations ne semble pas très intéressant.

- Enfin, pour  $(d_o, d_b) = (\mathcal{W}_2, \mathcal{H}^{-1})$  (cf. la figure 4.1f), le minimiseur aura tendance à se décaler de  $\rho_0$ , mais en même temps de croître au niveau de  $\rho_1$ . Il ne semble pas y avoir de correction de la position.

L'explication de ce résultat par les conditions d'Euler-Lagrange est difficile à obtenir puisque ces conditions ne donnent pas de formulation explicite de la valeur de  $\rho_0^a$ .

En résumé, dans cet exemple de 3D-Var, retrouver une analyse de position corrigée n'est possible qu'avec le choix  $d_o = d_b = \mathcal{W}_2$ .

Nous soutenons que lorsqu'un modèle est impliqué, modèle faisant évoluer  $\rho_0$  et donnant directement les observations  $\rho^{\text{obs}}$ , nous obtiendrons le même type de courbe, pour les mêmes raisons de condition d'Euler-Lagrange expliquées pour  $(d_o, d_b) = (\mathcal{H}^{-1}, \mathcal{L}^2)$  ci-dessus. Notamment, mixer les distances ne semble pas intéressant.

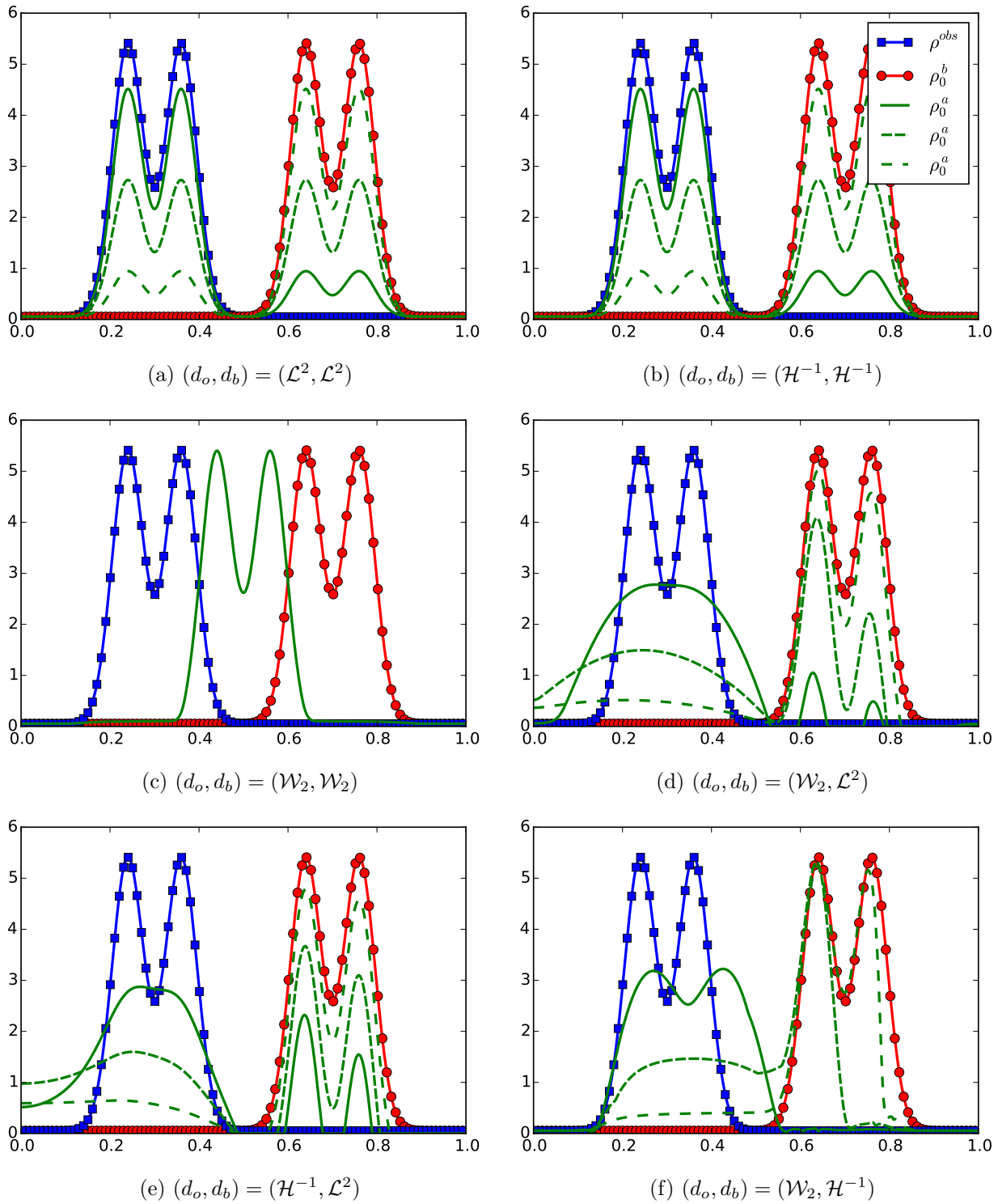


FIGURE 4.1 – Différentes analyses  $\rho_0^a$  de la minimisation de  $\mathcal{J}(\rho_0) = d_o(\rho_0, \rho_0^b)^2 + \omega_b d_b(\rho_0, \rho_1)^2$  pour plusieurs choix de couples  $(d_o, d_b)$  pris parmi les distance  $\mathcal{L}^2$ ,  $\mathcal{H}^{-1}$  et  $\mathcal{W}_2$ , et pour plusieurs valeurs de  $\omega_b$ . La valeur de  $\rho_0^b$  serait quelque part entre l'ébauche et l'observation.

## 4.2 Description des expériences suivantes

Dans ce chapitre nous réaliserons plusieurs expériences d'assimilation de données. Il n'y aura que  $\rho_0$  à contrôler, en assimilant des observations  $\rho_i^{\text{obs}}$  de  $\rho(t)$ . Les expériences seront de type jumelles (cf. la définition 3.2.1 sur ces expériences). Nous choisirons un modèle  $\mathcal{M}$  préservant l'appartenance à  $\mathcal{P}(\Omega)$ ,

$$\forall i, \quad \mathcal{M}_i : \mathcal{P}(\Omega) \rightarrow \mathcal{P}(\Omega)$$

avec  $\mathcal{M}_i$  intégrant une condition initiale  $\rho_0$  jusqu'au temps  $t_i$ . Les observations seront ainsi des éléments de  $\mathcal{P}(\Omega)$ . Elles seront dans un premier temps parfaites (créées sans bruit à partir d'un  $\rho_0^t$ ), puis nous ajouterons du bruit pour voir son influence.

Comme nous l'avons vu avant, si nous utilisons la distance de Wasserstein comme distance sur les observations, alors nous devons utiliser la même distance pour le terme d'ébauche, sans quoi le minimiseur ne sera pas réaliste. Ainsi, nous écrivons la fonction coût de Wasserstein  $\mathcal{J}_W$  :

$$\mathcal{J}_W(\rho_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{M}_i(\rho_0), \rho_i^{\text{obs}})^2 + \omega_b \mathcal{W}_2(\rho_0, \rho_0^b)^2$$

avec  $\rho_0, \rho_0^b, \rho_i^{\text{obs}} \in \mathcal{P}(\Omega)$ . Les résultats obtenus en minimisant  $\mathcal{J}_W$  seront comparés avec ceux obtenus par la fonction coût euclidienne donnée par :

$$\mathcal{J}_2(\rho_0) = \sum_{i=1}^{N^{\text{obs}}} \|\mathcal{M}_i(\rho_0) - \rho_i^{\text{obs}}\|_2^2 + \omega_b \|\rho_0 - \rho_0^b\|_2^2$$

ainsi qu'avec la fonction coût  $\mathcal{H}^{-1}$  :

$$\mathcal{J}_H(\rho_0) = \sum_{i=1}^{N^{\text{obs}}} \|\mathcal{M}_i(\rho_0) - \rho_i^{\text{obs}}\|_{-1}^2 + \omega_b \|\rho_0 - \rho_0^b\|_{-1}^2,$$

même si on s'attend à ce que cette fonction coût donne le même minimum que  $\mathcal{J}_2$ .

Pour minimiser les différentes fonctions coûts, nous utilisons une descente de plus forte pente. Pour la fonction coût  $\mathcal{J}_W$ , nous utilisons le gradient de Wasserstein et la minimisation expliquée dans la partie 3.3.6. Le critère d'arrêt sera basé sur la norme du gradient : dès que celui-ci devient plus petit que  $10^{-8}$ , l'algorithme s'arrête.

Le but des expériences suivantes sera :

- de voir à quoi ressemble le minimum dans le cas où quelques modèles sont impliqués, voir les sections 4.3, 4.4 et 4.5 ;
- d'étudier la capacité à récupérer les détails, voir la section 4.6 ;
- de comprendre l'influence des erreurs de position que contient l'ébauche ;
- d'observer l'influence sur le résultat des erreurs d'amplitude et de position que contiennent les observations, voir les sections 4.7 et 4.8.

### 4.3 Contrôle d'une seule densité dans un modèle linéaire.

Dans cette partie nous contrôlerons une densité  $\rho_0$  à partir de quelques observations et d'un modèle linéaire simple qui est une équation de continuité :

$$\mathbf{M}_{\rho_0}: \begin{cases} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) = 0 \\ \rho(0) = \rho_0 \end{cases}$$

avec  $u(t, x)$  le champ de vitesse indépendant de  $\rho$ , qu'on a pris égal à  $\sqrt{1-x}(1+0,2 \sin(2\pi x))$  qui fait translater la condition initiale vers la droite. Ce modèle est linéaire en  $\rho_0$  et conserve l'appartenance à  $\mathcal{P}(\Omega)$ .

L'ébauche, l'état vrai ainsi que les observations sont affichées sur la même figure 4.2a. Il n'y a pas d'erreurs sur les observations, seule l'ébauche contient une erreur de position. Il y a 5 observations et  $\omega_b = 4$ . On utilise les fonctions coût  $\mathcal{J}_{\mathcal{W}}$ ,  $\mathcal{J}_2$  ou  $\mathcal{J}_{\mathcal{H}}$  pour retrouver la condition initiale.

Après minimisation, on trace les analyses issues des différentes fonctions coûts. Les résultats issus des coûts  $\mathcal{J}_{\mathcal{H}}$  et  $\mathcal{J}_2$  sont équivalents et nous les affichons sur la figure 4.2b : pour ces deux distances l'analyse est une moyenne de l'ébauche et de l'état vrai. Alors que pour la distance de Wasserstein (voir figure 4.2c), l'analyse ressemble à un barycentre de Wasserstein entre l'ébauche et l'état vrai.

Nous pourrions penser que ces défauts initiaux (erreurs d'amplitude pour les analyses de  $\mathcal{J}_2$  et  $\mathcal{J}_{\mathcal{H}}$ , et erreurs de position pour l'analyse de  $\mathcal{J}_{\mathcal{W}}$ ) sont corrigés à la fin du modèle. Mais en traçant la sortie du modèle sur la figure 4.2d, nous constatons que ce n'est pas le cas. Ceci est dû au fait que le modèle est linéaire et qu'il n'y a qu'une variable à contrôler.

Cette première expérience montre que la minimisation des fonctions coûts a été réussie et donne des analyses ressemblant aux barycentres pour les distances utilisées. Nous avons notamment retrouvé une analyse de la forme souhaitée en utilisant la distance de Wasserstein.

Nous réalisons maintenant une expérience similaire utilisant le modèle non-linéaire de Saint-Venant.

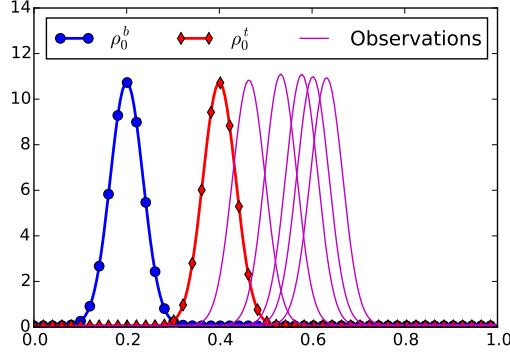
### 4.4 Assimilation dans un modèle de Saint-Venant

Le modèle de Saint-Venant, ou *Shallow-Water*, est un modèle décrivant l'évolution de la surface libre d'un fluide soumis à la gravité. Ce sont les équations de Navier-Stokes dans l'approximation où la hauteur du fluide est faible par rapport à la longueur du domaine, ce qui permet de faire quelques simplifications. Les équations de Saint-Venant sont très utilisées pour la modélisation océanographique. Le modèle 1D à fond plat est non-linéaire et s'écrit :

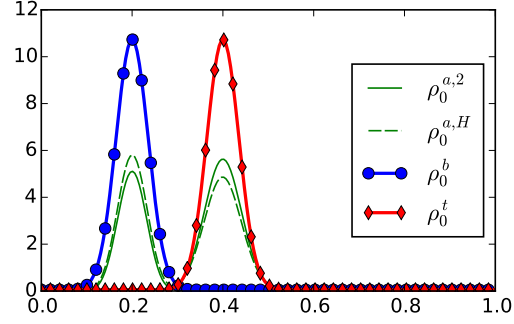
$$\mathcal{M}(h_0): \begin{cases} \frac{\partial h}{\partial t} + \operatorname{div}(hu) = 0 \\ \frac{\partial u}{\partial t} + u \cdot \nabla u + g \nabla h = 0 \\ h(0) = h_0, \quad u(0) = 0, \end{cases}$$

avec les conditions aux bords  $u(x=0) = u(x=1) = 0$  qui simulent une réflexion sur le bord. On ne contrôlera que  $h_0$  et pas la condition initiale sur  $u$  qui est fixée à 0.

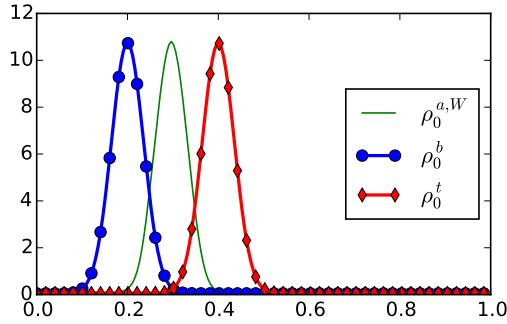
Le modèle préserve l'appartenance à  $\mathcal{P}(\Omega)$  puisque  $u$  est nul aux bords. Ainsi, si la variable  $h_0$  appartient à  $\mathcal{P}(\Omega)$ , alors à la sortie du modèle  $h(t)$  appartient aussi à  $\mathcal{P}(\Omega)$ . Nous ne notons plus la variable de contrôle  $\rho_0$  mais bien  $h_0$ , pour la hauteur de la surface libre.



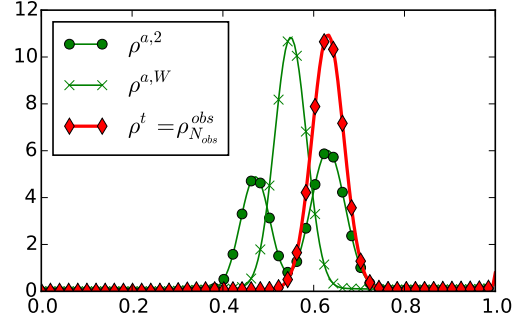
(a) L'ébauche, l'état vrai et les observations créées par intégration du modèle à partir de l'état vrai.



(b) Analyse  $\rho_0^{a,2}$  obtenue en utilisant le coût  $\mathcal{J}_2$ , qui ressemble à celle obtenue pour le coût  $\mathcal{J}_H$  ( $\rho_0^{a,H}$ ).



(c) Analyse  $\rho_0^{a,W}$  obtenue en utilisant la distance  $\mathcal{J}_W$ .



(d) Comparaison entre la dernière observation et les analyses  $\rho^{a,2}$  et  $\rho^{a,W}$  à la fin de la fenêtre d'assimilation (à la sortie du modèle).

FIGURE 4.2 – (a) Courbes d'ébauche, d'état vrai, et d'observations, comparées avec les analyses obtenues en utilisant les différentes distances  $\mathcal{L}^2$  et  $\mathcal{H}^{-1}$  (b) et  $\mathcal{W}_2$  (c), et leur valeur à la sortie du modèle en (d).

L'ébauche et l'état vrai sont des densités localisées et sont affichés sur la figure 4.3. Nous y avons également affiché cinq des  $N^{\text{obs}} = 10$  observations ( $h_i^{\text{obs}}$ ). Comme  $u_0 = 0$ , la condition initiale localisée aura tendance à se séparer en deux ondes, chacune allant de son côté. Les observations sont assimilées pour contrôler  $h_0$  via les fonctions coûts  $\mathcal{J}_W$  et  $\mathcal{J}_2$ .

Concernant la fonction coût  $\mathcal{J}_H$ , nous devons dire que nous n'avons pas réussi à la faire converger. Et ceci à cause de la remarque 3.4.2 qui dit que la valeur de la projection  $\mathcal{H}^{-1}$  n'est pas explicite. Ainsi, dans la suite du manuscrit, nous aurons tendance à ne pas l'utiliser.

Les états analysés sont présentés dans la figure 4.4. Dans la figure 4.4a nous traçons l'analyse  $h_0^{a,2}$  associée à  $\mathcal{J}_2$ , au temps initial et plus tard ( $h, u$ ) dans la fenêtre d'assimilation. En figure 4.4b nous traçons l'analyse  $h_0^{a,W}$  correspondant à  $\mathcal{J}_W$  et aussi ( $h, u$ ) dans la fenêtre d'assimilation.

Le premier élément à remarquer en regardant cette figure est la valeur de l'état analysé de  $\mathcal{J}_W$ . Il consiste simplement en un très grand pic. Nous pourrions penser que dans la formulation continue du problème, la solution serait un Dirac. Avoir un Dirac n'est pas un problème pour la distance de Wasserstein, puisqu'elle favorise plus une bonne position de la matière plutôt qu'une bonne forme, contrairement à la distance euclidienne. Ainsi, avoir un pic bien placé (proche de l'ébauche, notamment) est bien pour la distance de Wasserstein. De plus, comme le modèle de Saint-Venant

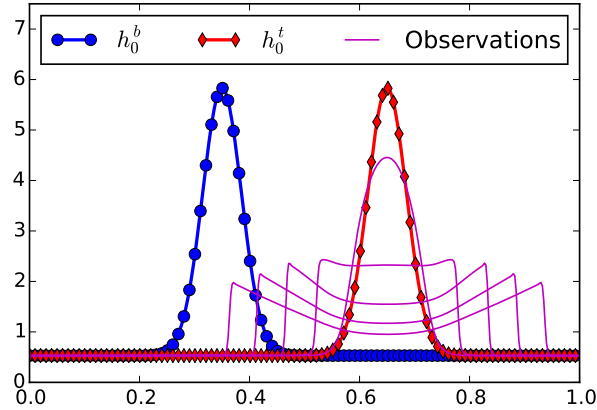


FIGURE 4.3 – Ébauche, état vrai et valeur de quelques observations utilisées pour contrôler la condition initiale de ce modèle de Shallow-Water.

est dispersif, ce pic est vite transformé en une onde ressemblant à celle de l'état vrai, comme nous pouvons le voir à droite dans la figure 4.4b.

Cette onde ressemble d'ailleurs plus à celle de l'état vrai (mais décalée en espace) que celle de l'analyse de  $\mathcal{J}_2$ . Cette dernière conserve la marque de l'ébauche et est donc constituée de quatre parties se déplaçant, que l'on peut distinguer dans la vitesse, cf. en bas à droite de la figure 4.4a.

Contre l'apparition du pic, nous proposons d'ajouter un terme de régularisation dans la fonction coût  $\mathcal{J}_{\mathcal{W}}$ .

## 4.5 Ajout d'un terme de régularisation

La fonction  $\mathcal{J}_{\mathcal{W}}$  ne pénalisant pas de fortes variations localisées, nous sommes amenés à ajouter un terme de régularisation pour lutter contre leurs créations. La fonction coût, s'écrit alors :

$$\mathcal{J}_{\mathcal{W}}(h_0) = \sum_i \mathcal{W}_2(\mathcal{M}_i(h_0), h_i^{\text{obs}})^2 + \omega_b \mathcal{W}_2(h_0, h_0^b)^2 + \mathcal{J}_r(h_0),$$

avec  $\mathcal{J}_r(h_0)$  pénalisant les pics. Nous pouvons prendre la norme euclidienne de  $h_0$  :

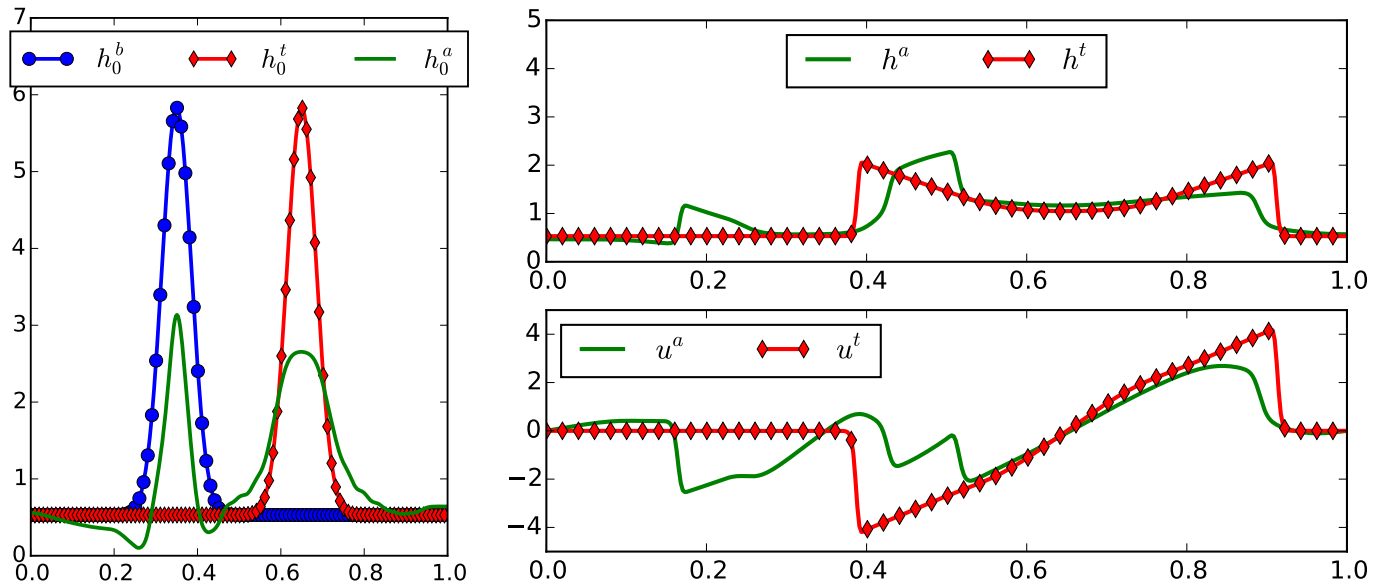
$$\mathcal{J}_r(h_0) = \alpha \|h_0\|_2^2 = \alpha \int_{\Omega} |h_0|^2 dx$$

qui pénalise les grosses valeurs, ou la norme  $\mathcal{H}^1$  :

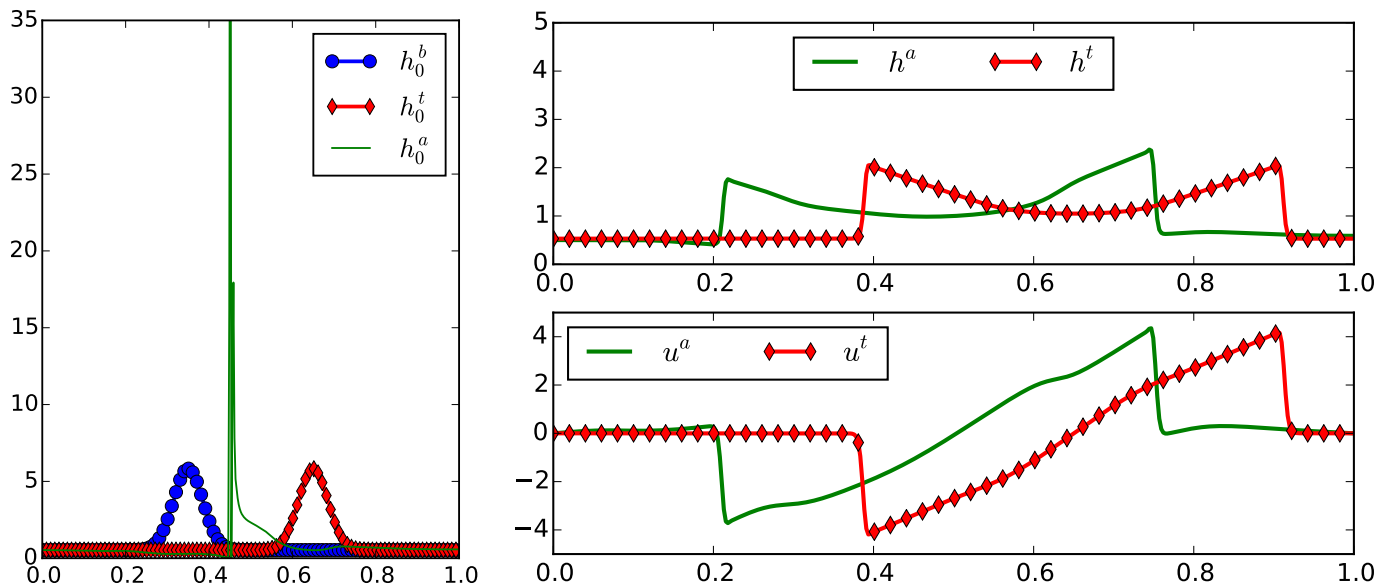
$$\mathcal{J}_r(h_0) = \beta \|\nabla h_0\|_2^2 = \beta \int_{\Omega} |\nabla h_0|^2 dx$$

qui pénalise les grosses variations.

Pour notre expérience, les deux régularisations réussissent à limiter la hauteur du pic. Pour la même expérience que dans la partie précédente, nous traçons sur la figure 4.5 les analyses en ayant ajouté une régularisation ou l'autre, et pour différentes valeurs des poids  $\alpha$  et de  $\beta$ . La hauteur du pic est limitée et il ne reste qu'une masse centrale. Il faut cependant faire attention à ne pas mettre le coefficient trop grand, sinon les données en sont trop affectées.



(a) À gauche, ébauche, état vrai et analyse correspondant à  $\mathcal{J}_2$ . À droite, état vrai et analyse plus tard dans la fenêtre d'assimilation. En haut à droite on affiche  $h(t)$  et en bas à droite on affiche la vitesse  $u(t)$ .



(b) Mêmes courbes que dans (a) mais avec l'analyse de  $\mathcal{J}_W$  à la place.

FIGURE 4.4 – Valeurs des analyses de  $\mathcal{J}_2$  et  $\mathcal{J}_W$ .

Nous remarquons aussi que le changement dans la fenêtre d'assimilation est minimale. En effet, nous traçons dans la figure 4.6 la valeur de  $h$  et  $u$  pour une des analyses ainsi régularisées (régularisation  $\dot{\mathcal{H}}^1$  avec  $\beta = 10^{-2}$ ), et nous ne remarquons que très peu de différence avec les courbes obtenues sans régularisation, tracées plus tôt dans 4.4b.

#### Remarque 4.5.1

L'analyse  $h_0^{a,W}$  obtenue dans la partie d'avant sans régularisation et affichée dans la figure 4.4b, même étant de forme non voulue, est légèrement meilleure selon le critère de la valeur  $\mathcal{J}_W$  (sans le terme de régularisation) que l'analyse régularisée (tracée dans la figure 4.6). En effet, la valeur de

$$\mathcal{J}_W(h_0) = \sum_i \mathcal{W}_2(\mathcal{M}_i(h_0), h_i^{\text{obs}})^2 + \omega_b \mathcal{W}_2(h_0, h_0^b)^2$$

vaut  $6.1 \cdot 10^{-3}$  pour la première et  $6.4 \cdot 10^{-3}$  pour la seconde, ce qui fait un écart relatif de 5,2%.

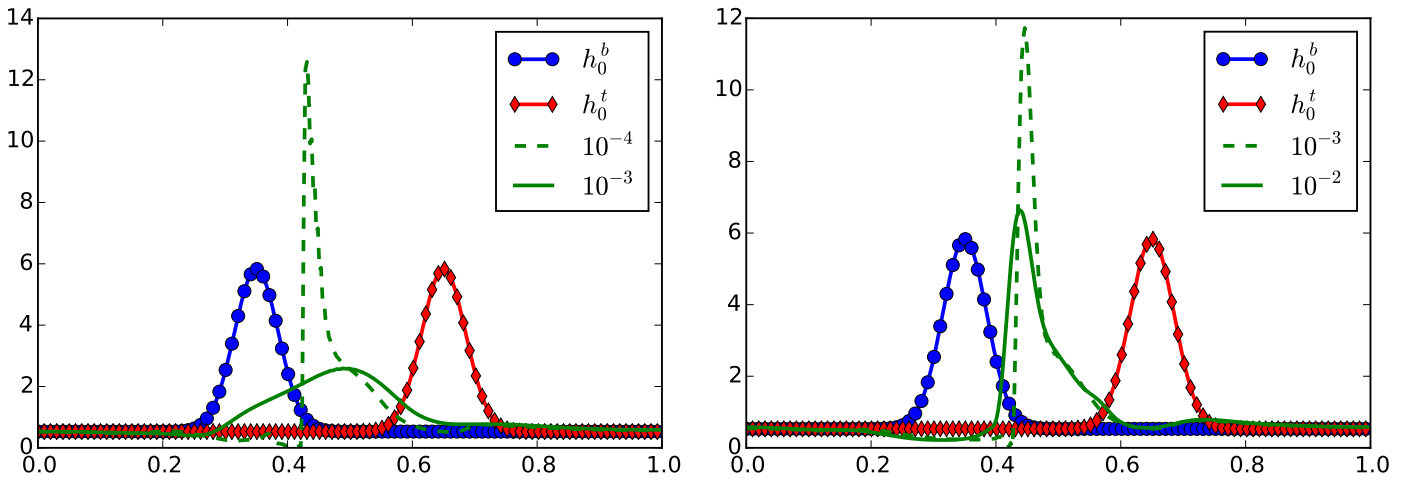


FIGURE 4.5 – L'état analysé correspondant à la fonction coût  $\mathcal{J}_W$  à laquelle on a rajouté une régularisation  $\mathcal{L}^2$  à gauche, et  $\dot{\mathcal{H}}^1$  à droite, pour différents poids de régularisation.

## 4.6 Capacité à retrouver la forme de l'objet

Nous avons utilisé des densités de forme simple, gaussienne, dont le barycentre pour la distance de Wasserstein est toujours une forme ressemblant à une gaussienne, même si un modèle est impliqué dans l'assimilation. Nous avons cependant dû rajouter des termes de régularisation à la fonction (pour éviter l'apparition de Diracs dans des modèles diffusifs ou dissipatifs).

Nous nous intéressons ici à savoir ce qu'il se passe lorsque la forme est plus complexe, c'est-à-dire avec des détails : est-ce que l'assimilation de données avec la distance de Wasserstein est capable de récupérer ces détails ? Un point faible de la distance de Wasserstein est qu'elle est insensible aux petites variations telles que ces détails. Nous pouvons nous attendre à ce que la distance de Wasserstein les récupère mal. C'est ce que nous étudions ici à travers un exemple.

Dans cette partie, nous réaliserons le même type d'expérience que précédemment, avec le modèle de Saint-Venant 1D. Cette fois-ci nous utilisons pour l'état vrai et l'ébauche des densités localisées mais contenant plusieurs pics, comme dans la figure 4.7a. L'intégration par le modèle Saint-Venant



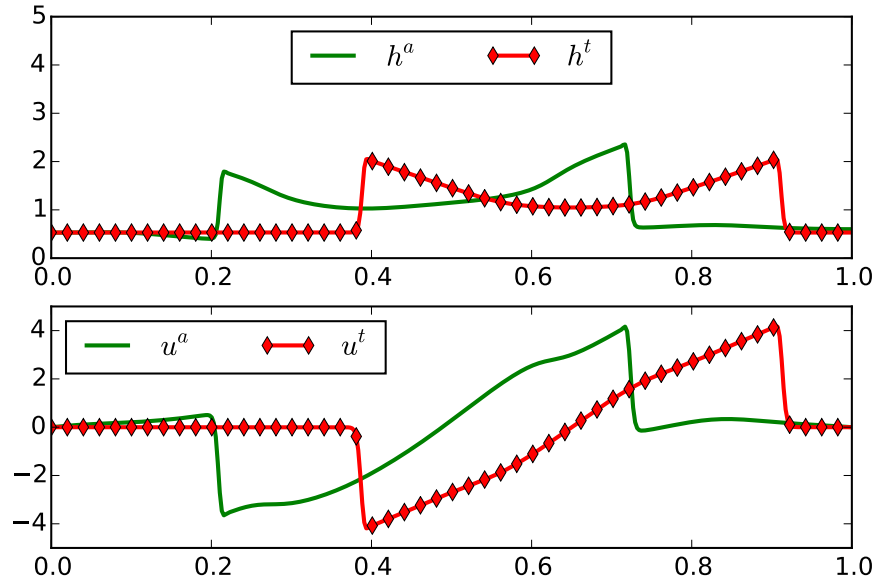
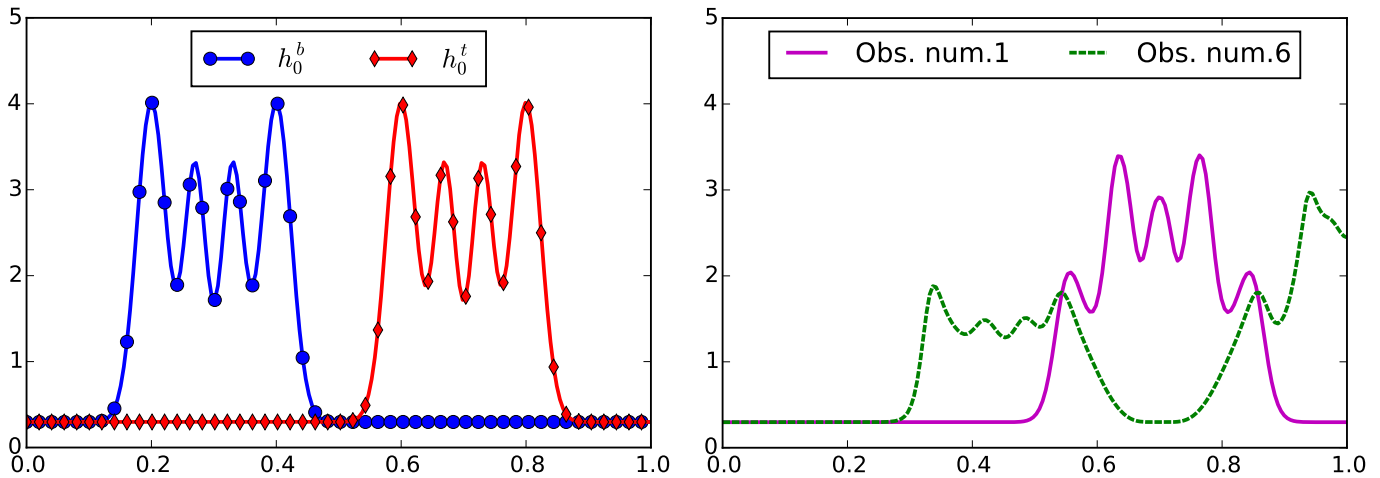


FIGURE 4.6 –  $h(t)$  et  $u(t)$  pour un temps  $t$  dans la fenêtre d'assimilation, pour l'ébauche, l'état vrai et l'analyse obtenue par une régularisation  $\mathcal{H}^1$  pour  $\beta = 10^{-2}$ .

a tendance à faire disparaître les détails par dispersion, mais ils sont quand même toujours présents comme on peut le voir dans la figure 4.7b où l'on affiche les première et sixième observations, qui sont des intégrations de  $h_0^t$  par le modèle. Il y a dix observations en tout.



(a) État vrai et ébauche utilisés pour l'expérience. Ici, ces états incluent des détails (des pics). (b) Première et sixième observations obtenues par intégration de  $\rho_0^t$  par le modèle.

FIGURE 4.7 – Données utilisées pour l'expérience de la section 4.6.

L'assimilation est réalisée en écrivant la fonction coût de Wasserstein classique, à laquelle il faut ajouter un terme de régularisation  $\mathcal{L}^2$  pour éviter l'apparition du Dirac. Nous ne mettons pas de régularisation  $\mathcal{H}^1$  puisqu'elle s'opposerait à l'apparition des détails. L'analyse est affichée figure 4.8.

On constate que l'analyse n'a pas retrouvé les détails au temps initial (donc également aux temps

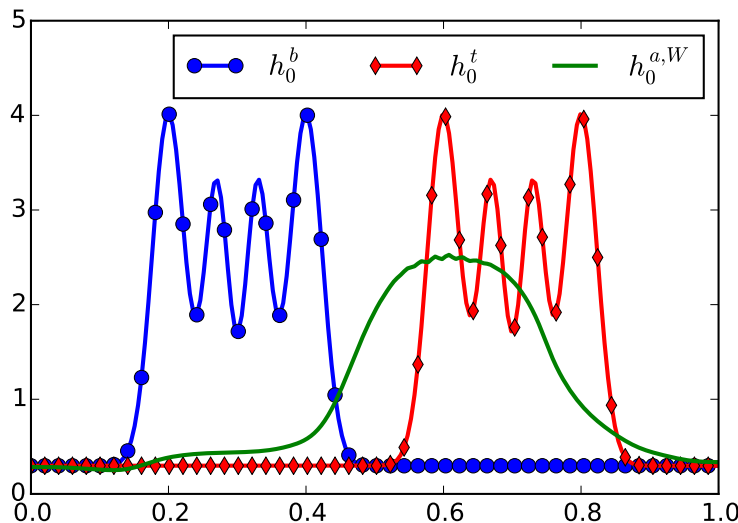
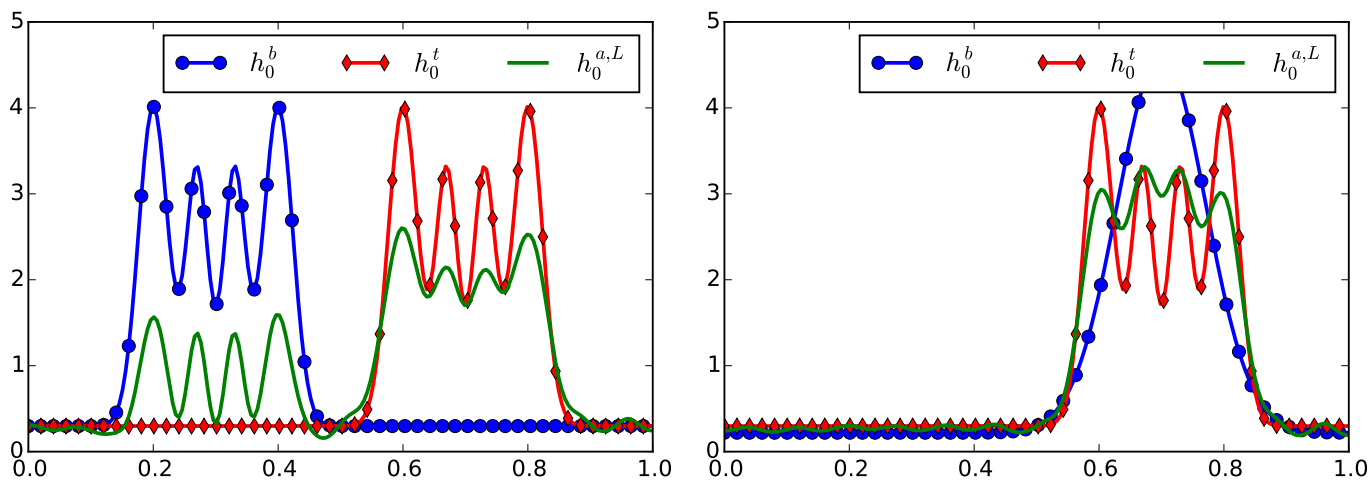


FIGURE 4.8 – Analyse pour la fonction coût  $\mathcal{J}_W$ . Celle-ci est placée au milieu, mais sans les détails des deux états utilisés.

supérieurs). Pourtant, la minimisation de la fonction coût va loin et elle s'est arrêtée pour la norme du gradient valant  $10^{-10}$ . Pour voir l'impact de la dispersion du modèle de Saint-Venant, nous avons mené la même expérience avec un coût euclidien classique, pour un critère d'arrêt équivalent et un terme d'ébauche égal. L'analyse pour la même expérience conduit à la figure 4.9a. Si l'on remplace l'ébauche où il y a de l'erreur de position par une autre contenant de l'erreur en amplitude, on obtient la figure 4.9b. Dans les deux cas, la fonction coût  $\mathcal{J}_2$  arrive à retrouver les détails.



(a) Analyse pour la fonction coût  $\mathcal{J}_2$  lorsque l'ébauche contient des erreurs de position. (b) Analyse pour la fonction  $\mathcal{J}_2$  lorsque l'ébauche a seulement des erreurs d'amplitude.

FIGURE 4.9 – La fonction  $\mathcal{J}_2$  arrive bien à retrouver les détails, même si le modèle les diffuse rapidement et sont rapidement invisibles dans les observations, et ce qu'on ait des erreurs de position (à gauche) ou d'amplitude sur l'ébauche (à droite).

Le fait que la distance de Wasserstein ne retrouve pas la forme de l'état vrai s'explique par sa non-sensibilité des détails. Par exemple, le terme de régularisation s'oppose fortement à la formation des détails. On le montre à travers un exemple simple de minimisation de

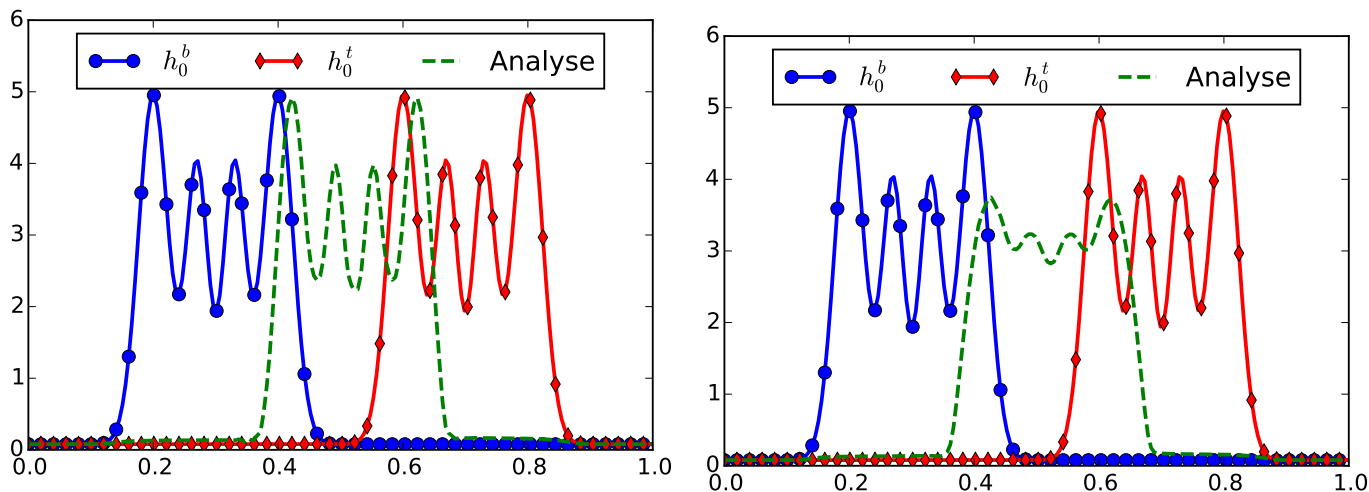
$$\mathcal{J}(h) = \mathcal{W}_2(h, h_0)^2 + \mathcal{W}_2(h, h_1)^2 + \alpha \|h\|_2^2. \quad (4.6.1)$$

Pour  $\alpha = 0$ , le minimum de  $\mathcal{J}$  se situe sur la géodésique reliant  $h_0$  et  $h_1$ . Les  $h_0$  et  $h_1$  sont des densités localisées présentant des détails. Sur la figure 4.10a nous montrons le minimum pour  $\alpha = 0$ , et sur la figure 4.10b le minimum pour  $\alpha = 10^{-4}$ , ce qui est insignifiant puisque

$$\frac{\alpha \|h\|_2^2}{\mathcal{J}(h)} \approx 1\%.$$

Pourtant, les détails sont bien atténués. Mais pour la distance de Wasserstein, ces deux analyses sont quasiment identiques puisque par exemple, la différence relative entre les deux analyses pour  $\mathcal{J}(h) = \mathcal{W}_2(h, h_0)^2 + \mathcal{W}_2(h, h_1)^2$  est de 0,03%.

Ainsi, en utilisant la distance de Wasserstein, il ne faudra pas s'attendre à pouvoir retrouver les détails de l'état vrai ou de l'ébauche, car ceux-ci seront fortement atténués par d'autres facteurs tels que le terme de régularisation.



(a) Analyse pour la fonction  $\mathcal{J}$  donnée en (4.6.1), sans régularisation ( $\alpha = 0$ ).

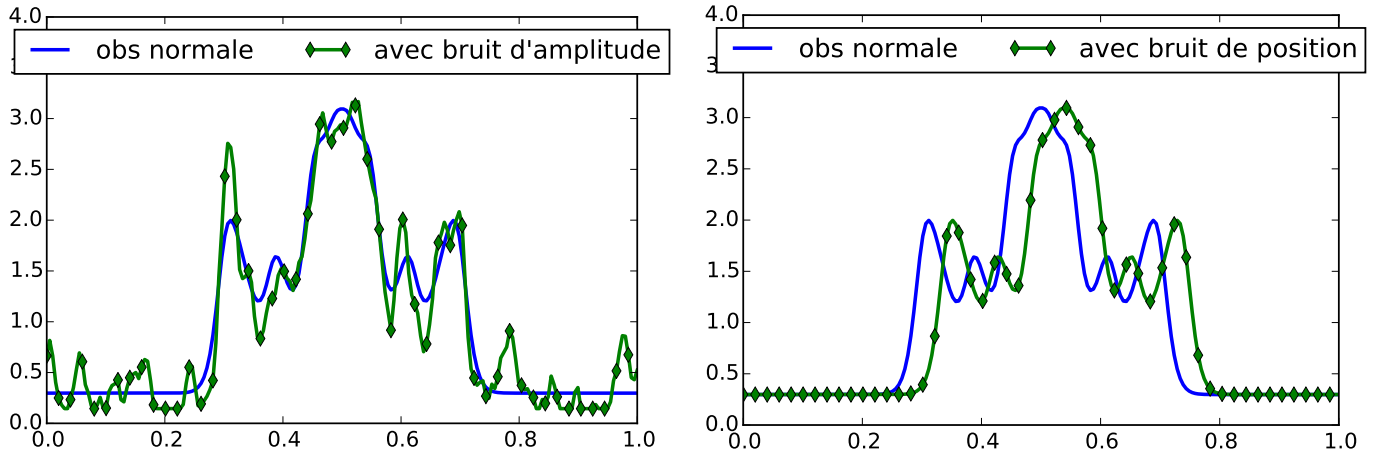
(b) Analyse pour  $\alpha = 10^{-4}$ .

FIGURE 4.10 – Analyses d'un exemple simple, sans régularisation à gauche (a), avec régularisation à droite (b).

## 4.7 Observations soumises à du bruit

Nous soumettons les observations à un bruit et l'on regarde son influence sur l'analyse, notamment lorsqu'on utilise une distance de Wasserstein. Cette fois-ci nous n'utilisons pas d'ébauche pour qu'elle n'influe pas sur le résultat. La régularisation sera d'abord enlevée. Le modèle utilisé est toujours le modèle de Saint-Venant.

Il y a deux types de bruit à ajouter sur les observations : soit un bruit d'amplitude (bruit blanc haute fréquence, cf. la figure 4.11a), soit un bruit de position (cf. la figure 4.11b).



(a) Observation à laquelle un bruit d'amplitude a été ajouté (b) Observation à laquelle un bruit de position a été ajouté

FIGURE 4.11 – Observations soumises à du bruit, d'amplitude à gauche (a), et de position à droite (b).

**Le bruit d'amplitude** On appelle bruit d'amplitude haute-fréquence un bruit additif, gaussien et de corrélation spatiale réduite. Ainsi, les observations sont

$$h_i^{\text{obs}} = \mathcal{M}_i(h_0^t) + \epsilon_i$$

avec  $\epsilon_i$  ce bruit. L'analyse est bien récupérée pour la fonction coût  $\mathcal{J}_2$  (cf. 4.12a). À l'inverse, l'analyse de Wasserstein (cf. 4.12b) est perturbée par ce bruit et l'a amplifié, pour que l'analyse colle aux observations. Nous rappelons que l'ajout d'un bruit haute-fréquence n'a que peu d'effet sur la valeur de  $\mathcal{J}_W$ . Par contre, lorsque nous ajoutons un petit terme de régularisation  $\mathcal{L}^2$  ( $\alpha = 10^{-4}$ ) à  $\mathcal{J}_W$ , le bruit sur l'analyse disparaît, mais aussi le recouvrement des détails.

**Le bruit de position** Nous ajoutons maintenant de l'erreur de position aux observations. Les erreurs de position sont faites ainsi. Au lieu de prendre les observations  $h_i^{\text{obs}}$  égales à  $\mathcal{M}_i(h_0^t)$ , nous poserons

$$h_i^{\text{obs}} := \mathcal{M}_i(\tau_i(h_0^t))$$

avec  $\tau_i$  un opérateur qui déplace spatialement  $h_0^t$  d'une constante  $d_i$  qui suit une loi de probabilité aléatoire de moyenne 0 et de variance 0,1, et indépendante entre deux observations.

Nous nous attendons à ce que le coût  $\mathcal{J}_W$  soit peu sensible à ces erreurs de position en retrouvant une analyse proche de la réalité, contrairement à l'analyse de  $\mathcal{J}_2$ .

En minimisant les fonctions coûts  $\mathcal{J}_2$  et  $\mathcal{J}_W$  nous obtenons les analyses respectivement présentées sur la figure 4.12c et 4.12d. Comme attendu, sous erreurs de position l'analyse pour  $\mathcal{J}_2$  est une sorte de diffusion de l'état vrai. L'analyse de  $\mathcal{J}_W$  est encore toute bruitée, mais ajoutant un petit terme de régularisation à  $\mathcal{J}_W$  nous remarquons que l'état a été retrouvé, aux détails près.

## 4.8 Dans le cas de présence d'erreur d'amplitude seulement

Nous avons vu que si l'ébauche et les observations ont toutes des erreurs de position, il valait mieux utiliser un coût utilisant les distances  $\mathcal{W}_2$ , et pour les termes d'écart à l'ébauche, et pour le terme d'écart aux observations. Mais que se passe-t-il si l'ébauche a des erreurs d'amplitude (et

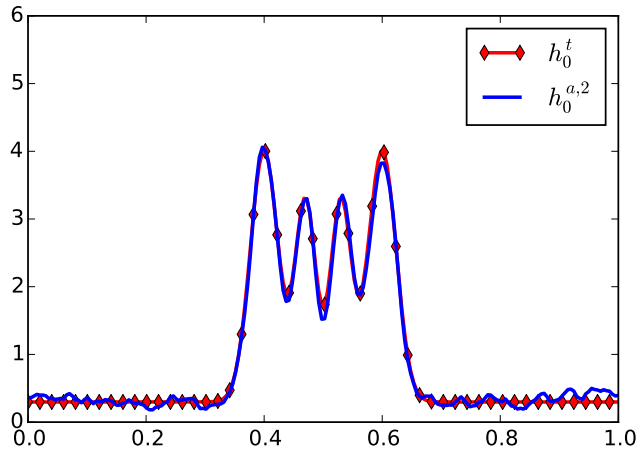
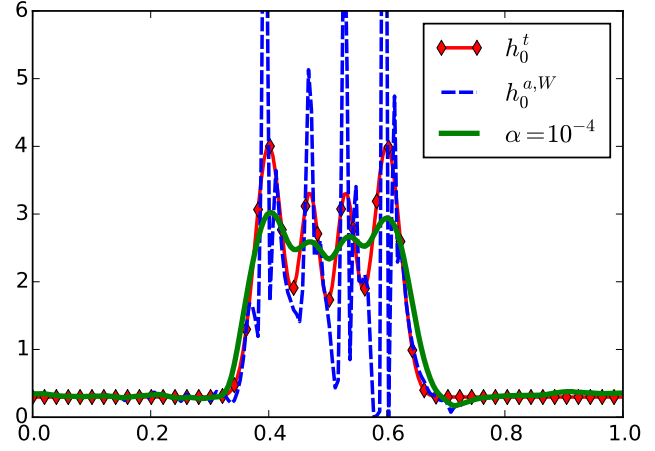
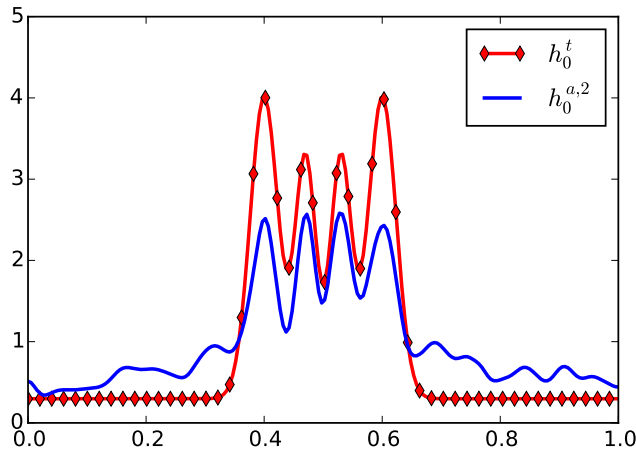
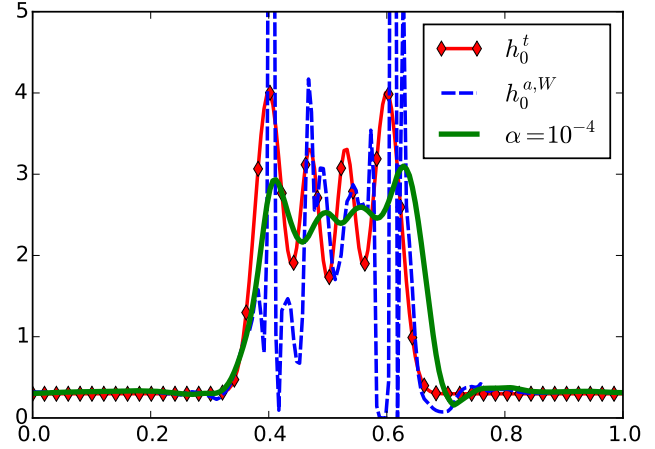
(a) Bruit d'amplitude : analyse pour un coût  $\mathcal{J}_2$ (b) Bruit d'amplitude : analyse pour un coût  $\mathcal{J}_W$ , sans et avec régularisation.(c) Bruit de position : analyse pour un coût  $\mathcal{J}_2$ (d) Bruit de position : analyse pour un coût  $\mathcal{J}_W$ , sans et avec régularisation.

FIGURE 4.12 – Analyses pour (a) un coût  $\mathcal{L}^2$  et (b) un coût  $\mathcal{W}_2$  lorsqu'il n'y a pas de bruit ajouté aux observations, lorsqu'il y a un bruit haute fréquence et lorsqu'il y a un bruit basse fréquence.

non de position), lorsque les observations ont des erreurs de position ? Dans ce cas, mesurer l'écart à l'ébauche par une distance de Wasserstein semble incompatible avec la définition du 4D-Var, où les erreurs d'amplitude impliquent des distances de Mahalanobis (distance euclidienne, cf. la démonstration des équivalences suivant la méthode 1.2.2).

Dans ce cas, nous écrivons alors la fonction coût comme :

$$\mathcal{J}(h_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{M}_i(h_0), h_i^{\text{obs}})^2 + \omega_h \|h_0 - h_0^b\|_2^2 \quad (4.8.1)$$

avec le terme d'écart aux observations prenant en compte les erreurs de position, et le terme d'écart à l'ébauche prenant en compte les erreurs d'amplitude. Nous effectuons du 3D-Var ( $\mathcal{M}_i = I$ ) pour tester cette fonction coût. Pour simuler des observations avec des erreurs de position, nous prenons pour les observations un décalage aléatoire de l'état vrai,  $h_i^{\text{obs}} = h_0^t(\cdot - d_i)$  avec  $d_i$  un nombre

aléatoire de moyenne nulle et de variance 0,05. L'ébauche sera une variation en amplitude de l'état vrai. Nous affichons sur la figure 4.13a ces états, deux observations décalées, ainsi que le minimum de  $\mathcal{J}$  pour un certain choix de  $\omega_h$ .

Nous avons plusieurs constats à faire :

- Nous avons bien une moyenne  $\mathcal{L}_2$  entre l'ébauche et l'état vrai, alors même que l'état vrai est observé avec des erreurs de position.
- Par contre, la convergence est lente : en effet, la descente utilisant le gradient euclidien est freinée par la distance de Wasserstein (et inversement). Il a fallu 249 itérations avant d'arriver à convergence (contre 5 itérations pour un coût tout Wasserstein).
- Un autre problème est si la moyenne des  $d_i$  n'est pas nulle : la moyenne de Wasserstein des observations sera décalée par rapport à l'ébauche, et donc nous aurons une analyse comme donnée dans la figure 4.1f : en deux parties et dont la forme diffère de celle des données.

Pour conclure cette discussion, réglant les deux derniers points et rejoignant le premier, nous pouvons remarquer que dans le même exemple, en mesurant l'écart à l'ébauche par une distance de Wasserstein, nous avons le même type de résultat, cf. figure 4.13b, mais plus rapidement (5 itérations). Il vaudra ainsi mieux utiliser une distance de Wasserstein si une partie des données seulement est sujette à des erreurs de position (dans le cas où une seule variable  $\rho_0$  est contrôlée, nous verrons d'autres exemples plus tard).

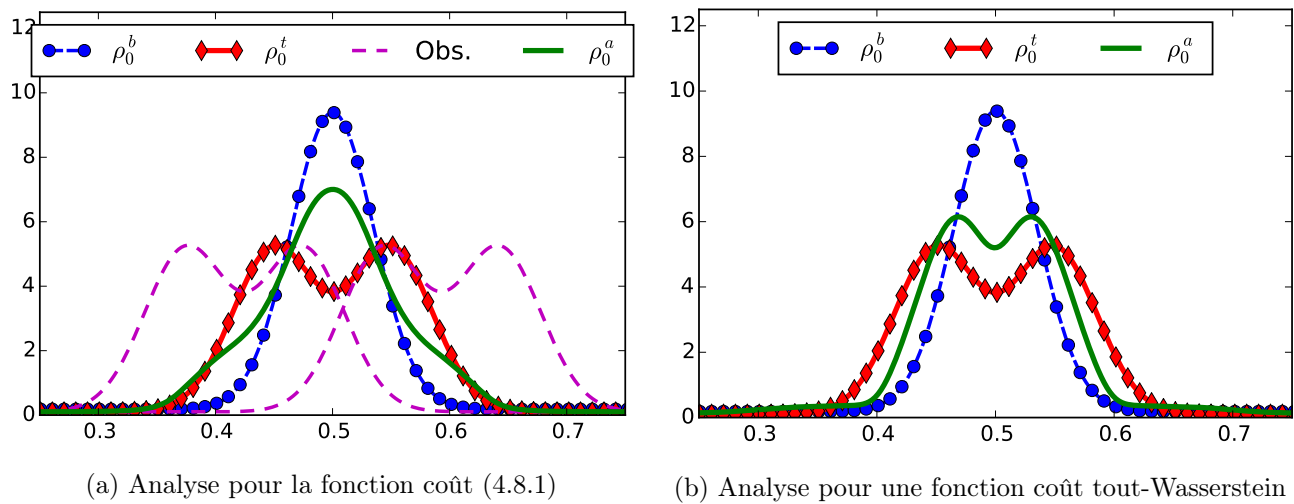


FIGURE 4.13 – Tracés de l'état vrai, de l'ébauche (avec erreur d'amplitude par rapport à l'état vrai) et des observations (avec erreurs de position par rapport à l'état vrai) utilisées pour l'expérience de la section 4.8.

## 4.9 Distance Wasserstein-Fisher-Rao pour des densités positives

La distance de Wasserstein a l'inconvénient de ne pouvoir calculer l'écart qu'entre deux densités de probabilité, ce qui limite son utilisation. Pour élargir le spectre d'applications du transport optimal pour l'assimilation de données, nous avons proposé aussi la distance Wasserstein-Fisher-Rao  $WF$  présentée dans la section 2.7 et qui calcule un transport optimal entre deux densités de masses positives.

Nous testons donc  $WF$ . Nous utilisons donc un modèle qui ne conserve que la positivité, et pas la masse. Ce modèle est une variation du modèle de Saint-Venant, nous avons rajouté un terme source :

$$\begin{cases} \partial_t h + \operatorname{div}(hu) = C\sqrt{h} \\ \partial_t u + u \cdot \nabla u + g\nabla h = 0. \end{cases}$$

Dans cette expérience, la masse de la condition initiale n'est pas prescrite. L'état vrai est de masse égale à 1, l'ébauche est de masse égale à 2. Les observations sont de masses supérieures à 1. Nous utilisons la distance  $WF$  pour assimiler les variables avec la fonction coût

$$\mathcal{J}_{\mathcal{W}}(h_0) = WF(\mathcal{M}(h_0), h^{\text{obs}})^2 + WF(h_0, h_0^b)^2.$$

Nous avons également ajouté des termes de régularisation  $\mathcal{H}^1$  pour éviter les oscillations et l'apparition de pic sur  $h_0$ . Les résultats sont similaires à ce que nous avons établi avec  $\mathcal{W}_2$  : nous trouvons une analyse placée entre l'ébauche et l'état vrai, de masse intermédiaire (à peu près égale à 1,3). La forme est légèrement altérée.

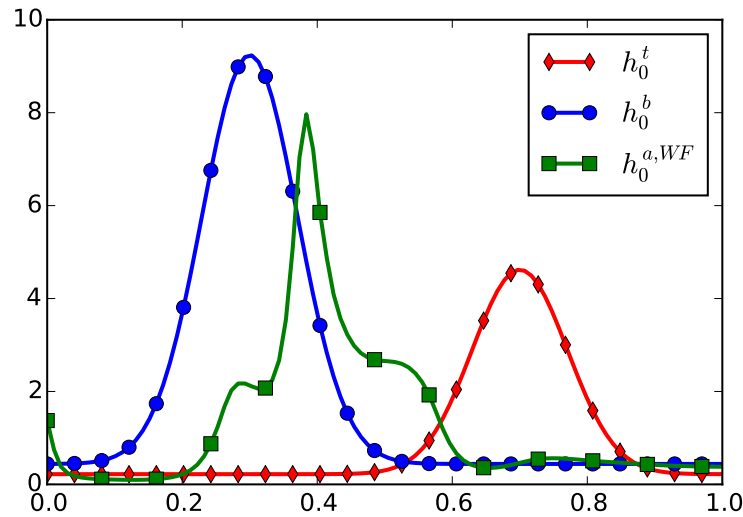


FIGURE 4.14 – Tracés de l'état vrai et de l'ébauche utilisés pour l'expérience d'assimilation à l'aide de la distance  $WF$ . L'analyse  $h_0^{a,WF}$  est également tracée.

Nous ne nous étendons pas plus sur cette distance parce qu'elle semble donner des résultats similaires à la distance de Wasserstein classique. Son principal problème dans le cadre de cette étude est le temps de son calcul, long, qui rendent les résultats délicats à obtenir.

## 4.10 Conclusion de ces contrôles d'une seule densité de probabilité

Nous avons réalisé plusieurs expériences d'assimilation de données pour contrôler une variable  $\rho_0 \in \mathcal{P}(\Omega)$  à l'aide de la distance de Wasserstein, ou  $\rho_0 \geq 0$  à l'aide de la distance  $WF$ . Ces données ont été prises localisées (c'est-à-dire que chaque donnée est non-nulle sur une zone réduite de l'espace).

Nous avons également utilisé la distance  $\mathcal{H}^{-1}$  pour écrire  $\mathcal{J}_{\mathcal{H}}$ , mais à cause du terme de projection, nous n'avons pas réussi à la faire converger. Étant donné que ses analyses semblent se rapprocher de celles de  $\mathcal{J}_2$ , nous avons décidé de ne plus l'utiliser.

Les caractéristiques de l'assimilation à l'aide de la fonction coût  $\mathcal{J}_{\mathcal{W}}$  sont les suivantes :

- Si l'ébauche et les observations possèdent des erreurs de position, alors  $\mathcal{J}_{\mathcal{W}}$  retrouve un élément  $\rho_0$  de position moyennée, corrigeant ainsi ces erreurs de position. Ce résultat s'oppose au résultat du coût classique  $\mathcal{L}^2$ , donnant un résultat étant une moyenne des données, et donc apparaît comme écrasée. La conséquence est que pour les temps supérieurs, l'analyse de  $\mathcal{J}_2$  n'est pas de la même forme que celle de l'ébauche et de l'état vrai, contrairement à l'analyse de  $\mathcal{J}_{\mathcal{W}}$ . Cela peut avoir des conséquences, si l'on pense par exemple à l'assimilation de données dans un scénario de cyclone, où les données (ébauche, état vrai) sont entachées d'erreurs de position :  $\mathcal{J}_2$  donnera deux cyclones de faible amplitude alors que  $\mathcal{J}_{\mathcal{W}}$  donnera un cyclone mieux placé et d'amplitude équivalente à ce qui a été observé.
- Il a tout de même fallu mettre un terme de régularisation  $\mathcal{H}^1$ , pour éviter les analyses de grande amplitude ou de forte variation, que la distance de Wasserstein seule ne pénalise pas. Cette régularisation a agi au temps initial en donnant une analyse d'amplitude réduite, mais pas aux temps supérieurs à cause de notre modèle dispersif de Saint-Venant.
- La forme des structures et les détails précis ne sont par contre pas bien retrouvés par la fonction coût  $\mathcal{J}_{\mathcal{W}}$ , n'étant pas importants pour la distance de Wasserstein.
- Sans terme de régularisation, les bruits d'amplitude impliquent une analyse  $\mathcal{J}_{\mathcal{W}}$  encore plus bruitée. Par contre, si nous ajoutons un terme de régularisation, tous ces bruits disparaissent.
- Nous pouvons éviter le mélange des distances  $\mathcal{W}_2/\mathcal{L}^2$  dans le cas où celles-ci sont justifiées par le fait qu'on ait des données entachées d'erreurs soit de position, soit d'amplitude. Le problème de ce mélange est qu'il peut donner des formes non souhaitées et en tout cas une minimisation très lente.

Ces résultats montrent donc que les erreurs de position peuvent être corrigés par une fonction coût  $\mathcal{J}_{\mathcal{W}}$ , mais cette correction est au prix en général des détails.





## Chapitre 5

# Assimilation dans un modèle complet de Saint-Venant

### Contenu du chapitre

5.1	Contrôle de la vitesse initiale . . . . .	111
5.1.1	Contrôle de la vitesse initiale seule et sans terme d'écart à l'ébauche . . . . .	111
5.1.2	Contrôles de la vitesse et de la densité initiales . . . . .	113
5.1.3	En présence d'erreurs de position . . . . .	116
5.2	Contrôle du fond du bassin . . . . .	117
5.2.1	Contrôle du fond seul, sans terme d'ébauche . . . . .	118
5.2.2	Contrôles du fond et de la hauteur de la colonne d'eau avec ébauche . . . . .	119
5.2.3	Utilisation d'une distance euclidienne pour le terme d'ébauche . . . . .	121
5.3	Contrôle de toutes les variables en même temps . . . . .	122
5.4	Conclusion générale de ces expériences . . . . .	125

Nous avons vu dans le chapitre 4 qu'une assimilation d'observations de type hauteur d'eau dans un modèle de Saint-Venant à l'aide de la distance de Wasserstein pouvait corriger la position de la condition initiale  $h_0$ . Nous allons maintenant tester la robustesse de la distance de Wasserstein au contrôle de plusieurs variables : aurons-nous encore une correction de la position ? Comment contrôler des variables qui ne sont pas des densités de probabilité ?

Pour répondre à ces questions nous allons mener des expériences de contrôle de plusieurs variables à la fois dans un modèle de Saint-Venant, comme précédemment, mais où une variable de fond  $f$  est ajoutée. Cette variable modélise un fond non plat, par exemple la présence d'obstacles perturbant la circulation de l'eau. Trois variables en tout sont impliquées dans ce modèle : la hauteur de la colonne d'eau  $h(t, x)$  de condition initiale  $h_0(x)$ , la vitesse  $u(t, x)$  de condition initiale  $u_0(x)$ , ainsi que la hauteur du fond  $f(x)$ , indépendante du temps, cf. figure 5.1. Nous avons ajouté cette variable pour rendre le problème plus réaliste.

Les équations du modèle Saint-Venant avec fond s'écrivent alors :

$$\mathcal{M}: \quad \begin{cases} \partial_t h + \operatorname{div}(hu) = 0 \\ \partial_t u + u \cdot \nabla u + g \nabla(h + f) = 0. \end{cases}$$

La vitesse sur le bord  $\partial\Omega$  vérifie  $u \cdot \vec{n} = 0$ . Nous devons distinguer la hauteur d'eau  $h$  et l'altitude de la surface libre,  $h + f$ . Par exemple, les observations  $(y_i^{\text{obs}})$  seront des captures de l'altitude de

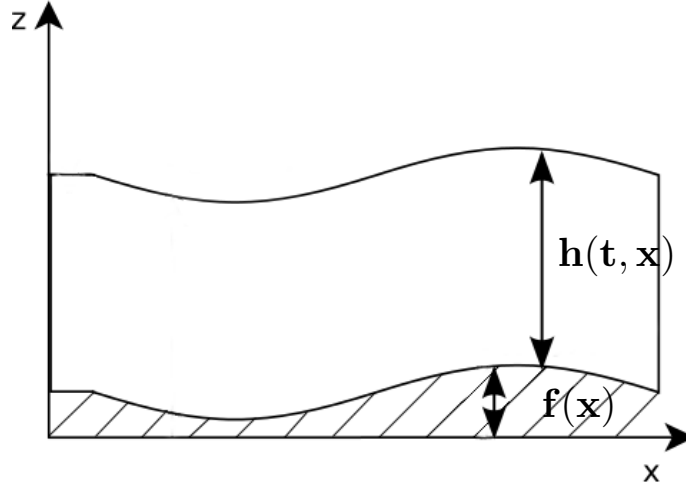


FIGURE 5.1 – Variables intervenant dans le modèle de Saint-Venant : au temps  $t$  et au point  $x$ ,  $h(t, x)$  est la hauteur de la colonne d'eau et  $f(x)$  est la hauteur du fond. La vitesse de l'eau  $u(t, x)$  n'est pas représentée ici.

la surface libre  $h(t) + f$  à des temps  $(t_i)$ . On définit l'opérateur général d'observation

$$\mathcal{G}_i : (h_0, u_0, f) \mapsto h(t_i) + f$$

avec  $h(t_i)$  la valeur de la variable  $h$  au temps  $t_i$ , selon le modèle  $\mathcal{M}$  et lorsque l'on prend  $h_0, u_0$  comme conditions initiales et  $f$  la hauteur du fond.

Comme dans le chapitre 4, nous supposons que  $h(t) \in \mathcal{P}(\Omega)$ . De plus,  $f$  sera considéré comme positif et de masse  $m_f$  fixée. Ainsi,  $f$  est un élément de  $m_f \mathcal{P}(\Omega)$  et une distance de Wasserstein pourra être utilisée pour comparer des fonds. Par contre, pour la vitesse  $u$ , nous n'imposons pas de contrainte de positivité ou sur la valeur de son intégrale et donc nous ne pourrions pas utiliser de distance de Wasserstein pour comparer deux vitesses.

Le but de ce chapitre est d'effectuer une assimilation d'observations  $y_i^{\text{obs}}$  en vue de contrôler les trois variables  $h_0, u_0$  et  $f$ . L'assimilation se fera à l'aide de la distance de Wasserstein sur les observations. Dans le chapitre 4, nous avons écrit un écart à l'ébauche avec la distance de Wasserstein : que se passe-t-il pour la vitesse? Nous étudierons ce point-ci en menant plusieurs expériences préliminaires avant de contrôler les trois variables à la fois.

1. D'abord, nous avons étudié le contrôle de la variable  $h_0$  seule dans le chapitre précédent 4.
2. Les contrôles de  $u_0$  seulement puis de  $h_0$  et  $u_0$  à la fois seront présentés dans la section 5.1.
3. Les contrôles de  $f$  seulement puis de  $h_0$  et  $f$  à la fois seront présentés dans la section 5.2.
4. Enfin, le contrôle de  $h_0, u_0$  et  $f$  à la fois seront présentés dans la section 5.3.

À l'issue de ces expériences nous aurons appris différentes leçons sur le fonctionnement de la fonction coût  $\mathcal{J}_{\mathcal{W}}$ , fonction de plusieurs variables.

Nous n'utiliserons dans ce chapitre ni la distance  $WF$  par simplicité, ni  $\mathcal{H}^{-1}$  pour la raison évoquée dans le chapitre précédent et dans la remarque 3.4.2.

## 5.1 Contrôle de la vitesse initiale

Nous allons effectuer une expérience d'assimilation pour pouvoir retrouver les  $h_0$  et  $u_0$  initiaux (le fond  $f$  ne sera pas contrôlé et sera posé égal à 0) en possession d'observations  $h_i^{\text{obs}}$  de  $h$  à des temps  $t_i > 0$ . L'assimilation variationnelle de données s'écrit :

$$\mathcal{J}_{\mathcal{W}}(h_0, u_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(h_0, u_0), h_i^{\text{obs}})^2 + d_h(h_0, h_0^b)^2 + d_u(u_0, u_0^b)^2.$$

Une question sera notamment de savoir quelle distance prendre pour les termes d'ébauche. Pour la distance à l'ébauche  $u_0^b$ , étant donné que la variable de vitesse n'est pas positive ou de masse fixée, définir une distance de Wasserstein entre  $u_0$  et  $u_0^b$  semble difficile. Nous allons d'abord étudier le cas où il n'y a pas de terme d'ébauche, avant d'en mettre un exprimé à l'aide de la distance  $\mathcal{L}^2$ .

Avant de contrôler  $h_0$  et  $u_0$  sans terme d'ébauche sur  $u_0$ , nous étudions le contrôle de  $u_0$  seul.

### 5.1.1 Contrôle de la vitesse initiale seule et sans terme d'écart à l'ébauche

Nous réalisons une expérience de contrôle de la vitesse initiale uniquement ( $f$  et  $h_0$  ne seront pas contrôlés et sont fixés à  $f^t$  et  $h_0^t$ ), à l'aide de la distance de Wasserstein, sans terme d'écart à l'ébauche. Pour retrouver  $u_0$ , nous utilisons la fonction coût :

$$\mathcal{J}_{\mathcal{W}}(u_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(u_0), h_i^{\text{obs}})^2.$$

Il n'y a pas de terme de régularisation. Il n'y a pas non plus d'erreurs dans les observations. Le but de cette expérience est double :

1. Identifier le produit scalaire à choisir pour calculer le gradient servant à minimiser la fonction. En effet, on a vu dans le chapitre 3 qu'il valait mieux utiliser le gradient de Wasserstein pour minimiser une fonction dépendante de  $h_0$  et écrite à l'aide de la distance de Wasserstein. Qu'en est-il ici, où le produit scalaire de Wasserstein n'est *a priori* pas défini pour des vitesses ?
2. Comparer la minimisation à celle de

$$\mathcal{J}_2(u_0) = \sum_{i=1}^{N^{\text{obs}}} \|\mathcal{G}_i(u_0) - h_i^{\text{obs}}\|_2^2.$$

**Établir le produit scalaire.** Nous allons choisir le produit scalaire à utiliser pour calculer le gradient servant dans la minimisation de la fonction  $\mathcal{J}_{\mathcal{W}}$  écrite à l'aide de la distance de Wasserstein. Si nous suivions l'approche du chapitre 3, nous établirions que le produit scalaire à choisir serait le produit scalaire de Wasserstein, qu'on pourrait définir pour  $u_0$  par

$$v = -\text{div}(u_0 \nabla \Psi), v' = -\text{div}(u_0 \nabla \Psi'), \langle v, v' \rangle = \int_{\Omega} u_0 \nabla \Psi \cdot \nabla \Psi'.$$

Mais ceci n'est pas un produit scalaire, puisque  $u_0$  peut être négatif.

C'est pourquoi nous ne choisirons pas ce produit scalaire pour calculer le gradient. Nous utiliserons plutôt le produit scalaire  $\mathcal{L}^2$  classique, pour le cas de la vitesse, pour la raison suivante : le gradient  $\mathcal{L}^2$  de  $\mathcal{J}$  est une vitesse multipliée par  $h$ . Dans le chapitre 3 nous avons vu qu'il valait mieux utiliser le gradient de Wasserstein parce qu'il était consistant avec les données. C'est le même constat que nous faisons ici. Décrivons le calcul du gradient  $\mathcal{L}^2$  à travers la proposition suivante :

**Proposition 5.1.1**

Pour une perturbation  $v$  de la vitesse  $u_0$ , on a :

$$\mathcal{J}_{\mathcal{W}}(u_0 + \varepsilon v) = \mathcal{J}_{\mathcal{W}}(u_0) + \varepsilon \langle v, \Psi \rangle_2 + \mathcal{O}(\varepsilon) \quad (5.1.1)$$

avec  $\Psi$  étant égal à

$$\Psi := \sum_i \mathbf{G}_i^*[u_0] \cdot \Phi_i^{\text{obs}}$$

avec, pour  $i \in \{1, \dots, N_{\text{obs}}\}$ ,  $\Phi_i^{\text{obs}}$  le potentiel de Kantorovich du transport optimal entre  $\mathcal{M}_i(u_0)$  et  $h_i^{\text{obs}}$ . Le  $\mathbf{G}_i^*$  est l'opérateur adjoint à  $\mathcal{G}_i$  et est défini par le modèle adjoint  $\mathbf{M}_i^*[u_0]$  de système d'équations :

$$\mathbf{M}_i^*[u_0]: \begin{cases} \partial_t \Phi_i + u \cdot \nabla \Phi_i + g \nabla \Psi_i = 0 \\ \partial_t \Psi_i + \text{div}(\Psi_i)u + h \nabla \Phi_i = 0 \\ \Phi_i(t_i) = \Phi_i^{\text{obs}}, \quad \Psi_i(t_i) = 0. \end{cases} \quad (5.1.2)$$

Alors,  $\mathbf{G}_i^*[u_0] \cdot \Phi_i^{\text{obs}} := \Psi_i(0)$ .

**Démonstration :** La preuve de la linéarisation (5.1.1) est similaire aux précédentes linéarisations de  $\mathcal{J}_{\mathcal{W}}$ . Pour la formule du système adjoint, on réfère le lecteur au théorème B.3.1. ■

Ainsi, le gradient  $\mathcal{L}^2$  de  $\mathcal{J}_{\mathcal{W}}$  est  $\Psi$ . On peut regarder quelle est la nature de  $\Psi$ , en observant son équation dans la deuxième ligne de (5.1.2). Nous remarquons que  $\Psi$  est homogène à  $h \nabla \Phi$ . Or  $\Phi$  est homogène à un potentiel de Kantorovich donc  $\nabla \Phi$  est une vitesse, et donc  $\Psi$  est une vitesse multipliée par  $h$ .

L'interprétation de ce résultat est la suivante : en faisant varier la vitesse initiale  $u_0$  en amplitude (variation spécifique au produit scalaire  $\mathcal{L}^2$ ), on joue en fait sur les déplacements de  $h(t)$  (variation spécifique à la distance de Wasserstein). Nous pourrions traduire ceci par : des erreurs d'amplitude sur  $u$  impliquent des erreurs de position sur  $h$ .

**Expérience et résultats** Pour la première expérience, nous utilisons le modèle de Saint-Venant à fond plat ( $f = 0$ ). Nous avons choisi  $h_0^t$ ,  $u_0^t$  et des observations comme disposés sur la figure 5.2. Le  $h_0^t$  est comme d'habitude une masse d'eau prête à se séparer en deux ondes. La condition  $u_0^t$  sera positive et nulle au bord (cf. figure 5.2a). On introduit aussi un détail dans  $u_0^t$  pour voir la capacité au processus d'assimilation à le retrouver. Étant donné que la vitesse initiale est positive, la masse aura tendance à se déplacer initialement vers la droite, comme on peut le voir à travers les observations tracées dans la figure 5.2b. Il y a en tout 20 observations et pour l'initialisation de la minimisation on choisit  $u_0^0 = 0$ . On arrête la minimisation après 30 itérations de descente de plus forte pente.

Nous minimisons la fonction  $\mathcal{J}_{\mathcal{W}}$  avec son gradient euclidien et la fonction  $\mathcal{J}_2$  également avec son gradient euclidien. Le minimum est 0 et est atteint en  $u_0^t$ .

On notera par l'exposant  $^W$  l'analyse trouvée par  $\mathcal{J}_{\mathcal{W}}$  et  $^2$  celle trouvée par  $\mathcal{J}_2$ . On affiche sur la figure 5.3a les analyses  $u_0^{a,W}$  et  $u_0^{a,2}$ .

Les résultats de convergence décrits ci-dessous sont encourageants.

1. L'analyse de  $\mathcal{J}_{\mathcal{W}}$  a réussi à trouver des détails de l'état vrai, comme vers  $x = 0,3$  où il y a une irrégularité dans  $u_0^t$ . Le détail est tout de même moins bien retrouvé que par  $\mathcal{J}_2$ .
2. La minimisation de  $\mathcal{J}_{\mathcal{W}}(u_0)$  a l'air d'avoir été performante, notamment grâce au choix du produit scalaire. On le voit à travers plusieurs critères. D'abord, les figures 5.3b et 5.3c

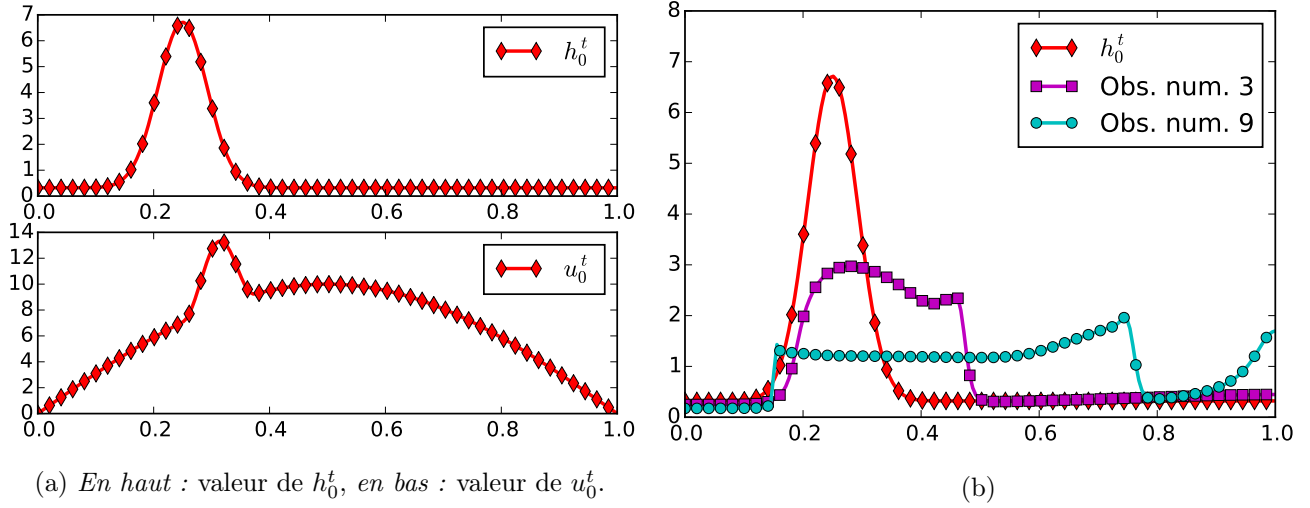


FIGURE 5.2 – Éléments utilisés pour la première expérience de contrôle de  $u_0$  sans ébauche :  $h_0^t$  en haut de (a),  $u_0^t$  en bas de (a) et quelques observations en (b).

montrent que  $u_0^{n,W}$  minimise mieux que  $u_0^{n,2}$  pour les critères de la valeur de  $\mathcal{J}_W$  et de la distance  $\mathcal{L}^2$  avec la solution vraie  $u_0^t$ . Sur la figure 5.3d cependant, on constate que l'analyse  $u_0^{a,2}$  se comporte un peu mieux pour  $\mathcal{J}_2$ .

3. Nous concluons qu'un coût écrit à l'aide de la distance de Wasserstein peut donc également retrouver une condition initiale ou un état qui est de type vitesse. Ce fait se confirme sur d'autres exemples (non exposés ici), où d'autres  $h_0^t$  et d'autres  $u_0^t$  ont été choisis.

Nous pouvons expliquer ces résultats de deux façons, qui sont probablement liées.

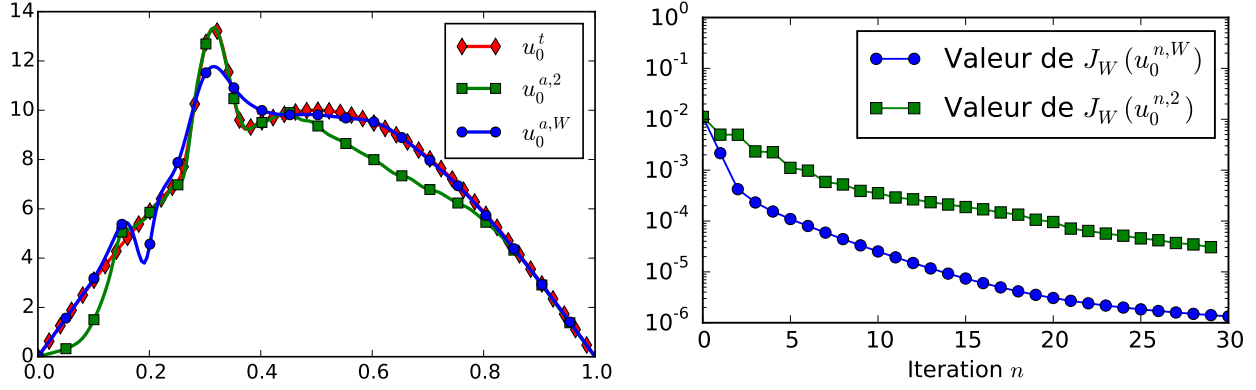
1. On rappelle l'article [Métivier et al., 2016] dont les résultats sont présentés dans la section 3.1. Dans cet article, les auteurs utilisent la fonction coût  $\mathcal{J}_W$  qui semble plus convexe que la fonction coût  $\mathcal{J}_2$  lorsque les sorties subissent un déplacement lors d'un changement des conditions initiales. Si c'est effectivement le cas ici, la minimisation est alors plus efficace.
2. La distance de Wasserstein est déjà la résolution d'un problème où il faut trouver la bonne vitesse, celle qui transporte l'observation calculée  $\mathcal{G}_i(u_0)$  sur l'observation  $h_i^{\text{obs}}$ . Ici, nous avons injecté le résultat de ce calcul sur  $u_0$ , après une intégration par le modèle adjoint. Donc, le calcul de la distance de Wasserstein est déjà un début de résolution du problème initial qui est de retrouver  $u_0^t$ .

Nous avons vu que l'on pouvait minimiser  $\mathcal{J}_W(u_0)$  à l'aide d'un gradient  $\mathcal{L}^2$ . Ceci nous conforte dans l'idée que nous pouvons choisir une ébauche écrite avec la distance  $\mathcal{L}^2$ . Ceci nous amène aussi à l'étape suivante qui est le contrôle simultané de  $u_0$  et  $h_0$ , en prenant également en compte des termes de rappel à l'ébauche et de régularisation.

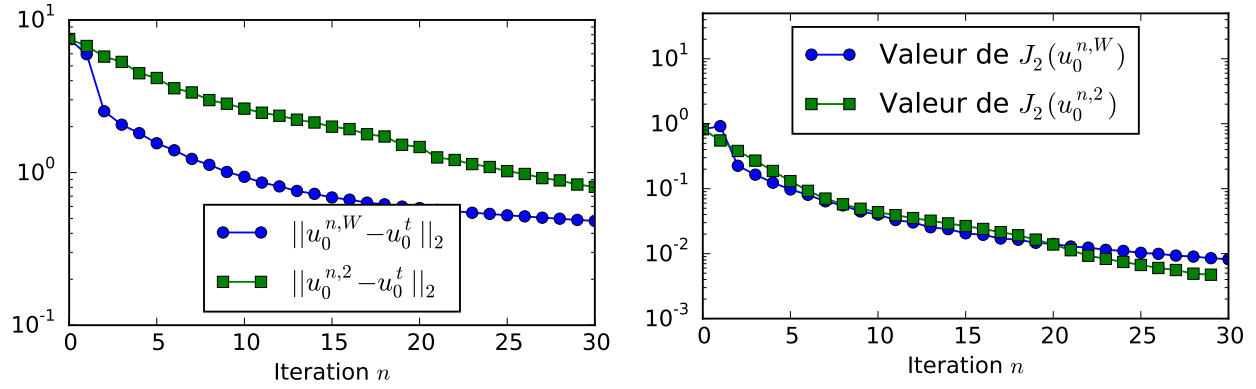
### 5.1.2 Contrôles de la vitesse et de la densité initiales

Le but est dans cette section de montrer la possibilité de pouvoir contrôler plusieurs données en même temps avec une fonction coût  $\mathcal{J}_W$  que l'on écrirait

$$\mathcal{J}_W(h_0, u_0) = \sum_i \mathcal{W}_2(\mathcal{G}_i(h_0, u_0), h_i^{\text{obs}})^2 + \omega_h \mathcal{W}_2(h_0, h_0^b) + \omega_u \|u_0 - u_0^b\|_2^2 + \mathcal{J}^r(h_0).$$



(a) Différentes analyses  $u_0^{a,2}$  et  $u_0^{a,W}$  comparées avec l'état vrai  $u_0^t$ . (b) Valeur de  $\mathcal{J}_W$  en fonction de l'itération  $n$  et des analyses.



(c) Valeur de  $\|u_0^n - u_0^t\|_2$  en fonction de l'itération  $n$  et des analyses. (d) Valeur de  $\mathcal{J}_2$  en fonction de l'itération  $n$  et des analyses.

FIGURE 5.3 – Résultats issus des minimisation de  $\mathcal{J}_W$  et  $\mathcal{J}_2$ , après 30 itérations de chacune des fonctions coût avec le même produit scalaire  $\mathcal{L}^2$ .

Nous avons ajouté un terme  $\mathcal{J}^r$  de régularisation  $\mathcal{H}^1$  sur  $h_0$  pour éviter les trop grandes valeurs et variations non pénalisées par les distances de Wasserstein. Nous comparerons les résultats trouvés par  $\mathcal{J}_W$  à l'analyse trouvée par un coût euclidien :

$$\mathcal{J}_2(h_0, u_0) = \sum_i \|\mathcal{G}_i(h_0, u_0) - h_i^{\text{obs}}\|_2^2 + \omega_h \|h_0 - h_0^b\|_2^2 + \omega_u \|u_0 - u_0^b\|_2^2.$$

L'ébauche  $h_0^b$  contiendra des erreurs de position.

Afin de comprendre l'influence du terme d'ébauche sur  $u_0$  sur le résultat de l'assimilation, nous allons effectuer une étape préliminaire où une assimilation est effectuée sans ce terme ( $\omega_u = 0$ ).

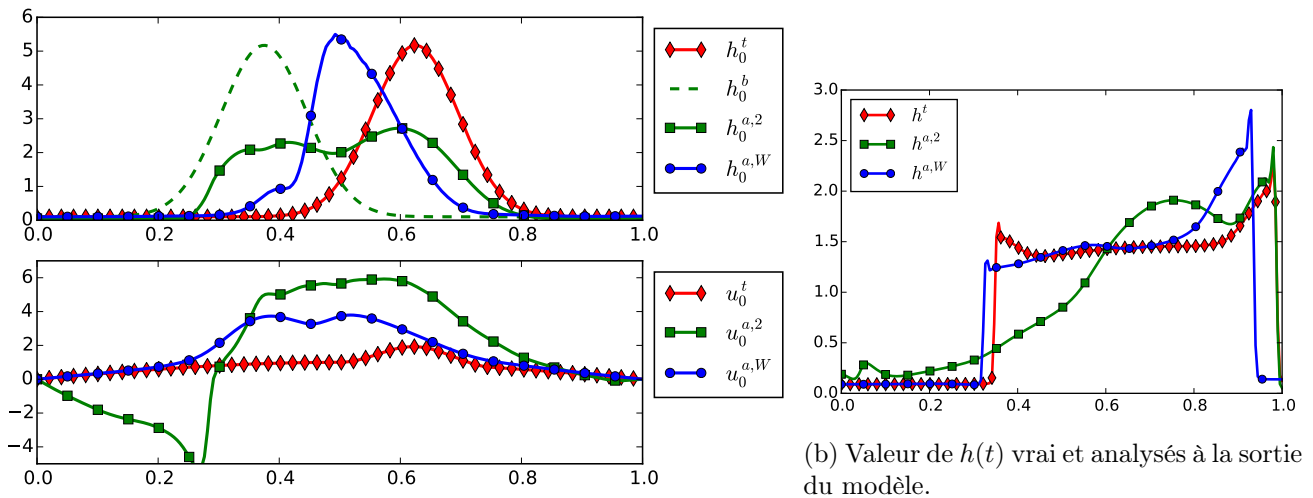
**Sans terme d'ébauche sur  $u_0$  ( $\omega_u = 0$ ).** Dans la fonction coût utilisée ici, il n'y a qu'un terme d'ébauche sur  $h_0$ . On s'attend à trouver un  $h_0$  entre l'ébauche et l'état vrai, comme dans le chapitre 4. Concernant la vitesse  $u_0$ , puisqu'elle est libre, on s'attend à ce qu'elle cherche à corriger le défaut de position entre l'analyse et l'état vrai de telle sorte que l'analyse colle aux observations.

On trace sur la figure 5.4a les valeurs de  $h_0^t$ ,  $h_0^b$ ,  $u_0^t$  utilisés pour cette expérience, ainsi que l'analyse trouvée par  $\mathcal{J}_W$ , que l'on compare avec celle trouvée par  $\mathcal{J}_2$ .

Concernant  $\mathcal{J}_W$ , on trouve en effet que l'analyse  $h_0^{a,W}$  est une masse placée entre l'ébauche et l'état vrai. La vitesse est surestimée, cela a pour conséquence de pousser initialement la masse vers la droite et de corriger son défaut de position pour pouvoir mieux coller aux observations. On peut le constater sur la figure 5.4b, qui affiche  $h^t(t)$  et  $h^{a,W}(t)$  à  $t > 0$ , l'analyse se retrouve assez proche en position de l'état vrai.

On peut comparer avec l'analyse de  $\mathcal{J}_2$  qui est également affichée sur la figure 5.4a. La densité  $h_0^{a,2}$  est une moyenne entre l'ébauche et l'état vrai. Elle est donc plus étalée, ce qui va faire que la vitesse initiale va être beaucoup plus élevée pour ramener toute la masse au niveau de l'état vrai et des observations, comme nous pouvons le voir sur la figure 5.4b.

Concernant la forme de la solution  $h(t)$  à la sortie du modèle, celle de  $\mathcal{J}_W$  est plus proche de la réalité que celle de  $\mathcal{J}_2$ , beaucoup plus étalée.



(a) En haut : Valeurs des différents  $h_0$  vrai, d'ébauche, et analysés.  
En bas : De même pour la vitesse  $u_0$ .

FIGURE 5.4 – Différentes fonctions  $h$  et  $u$  obtenues pour cette assimilation ayant pour but de contrôler  $h_0$  et  $u_0$ , dans le cas où il n'y a pas d'ébauche sur  $u_0$ .

**Avec un terme d'ébauche sur  $u_0$ .** On rajoute cette fois-ci un terme d'ébauche pour évaluer son influence. On ne change pas le reste des données de l'expérience. Normalement, ce terme d'ébauche limitera les surestimations de la vitesse, ce qui aura pour conséquence de limiter la correction de position apportée par la vitesse. On peut s'attendre à ce que les masses d'eau  $h_0$  soient mieux placées.

Nous traçons sur la figure 5.5a les états vrais et d'ébauche utilisés (dont  $u_0^b$ , inférieur à  $u_0^t$ ). Nous y avons mis aussi les analyses issues de  $\mathcal{J}_W$  et  $\mathcal{J}_2$ . Nous comparons ces résultats avec ceux donnés sans le terme d'ébauche, décrits dans le paragraphe précédent.

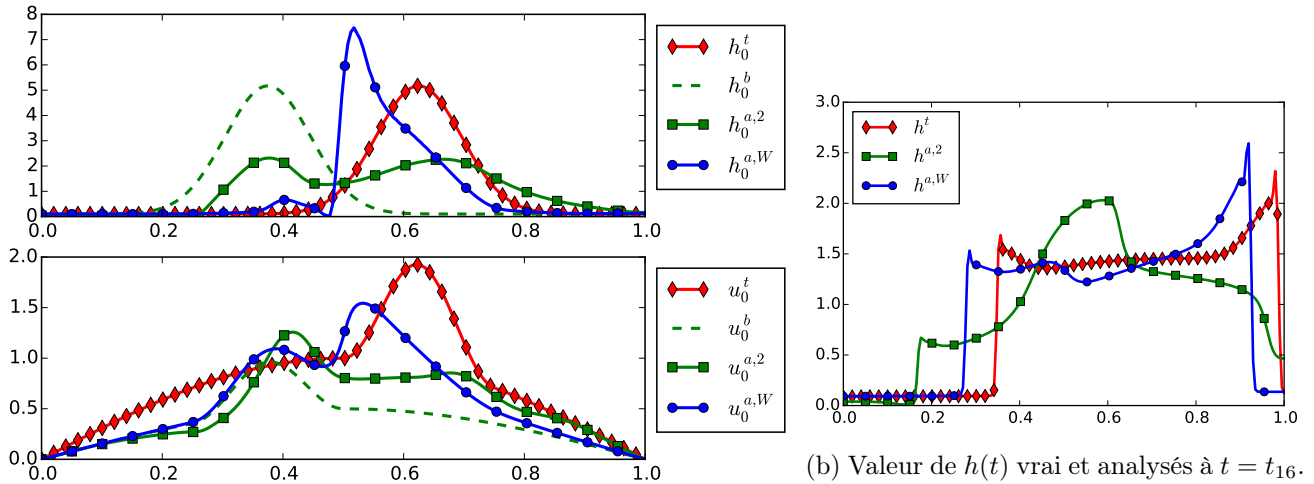
La vitesse est maintenant plus proche de l'ébauche, tout en cherchant à être la plus grande possible, pour coller aux observations. Le fait que la vitesse soit plus faible a pour effet de limiter la correction de position par rapport aux observations que nous avons vu avant. Mais cela n'a pas affecté la position de l'analyse  $h_0^{a,W}$  : nous voyons qu'elle n'a pas été repositionnée vers l'état vrai, et aux temps supérieurs est moins proche en position des observations (cf. figure 5.5b).

L'analyse  $h_0^{a,2}$  de  $\mathcal{J}_2$  est toujours une moyenne entre l'état vrai et l'ébauche. Étant donné que la vitesse est également limitée, nous constatons que l'analyse  $h^{a,2}$  aux temps supérieurs (cf. figure



5.5b) est plus loin des observations.

Concernant la forme qualitative de la solution à la fin de la fenêtre d'assimilation (cf. figure 5.5b), nous observons que l'analyse pour  $\mathcal{J}_W$  est toujours de même forme que l'état vrai, alors que l'analyse pour  $\mathcal{J}_2$  est toujours étalée.



(a) *En haut* : Valeurs des différents  $h_0$  vrai, d'ébauche, et analysés.  
*En bas* : De même pour la vitesse  $u_0$ .

FIGURE 5.5 – Valeurs des différentes fonctions pour cette fois-ci l'assimilation avec un terme d'ébauche sur  $u$ . On compare cette figure avec la figure 5.4 pour évaluer l'impact du terme d'ébauche.

### 5.1.3 En présence d'erreurs de position

Nous soumettons maintenant les 20 observations à des erreurs de position pour évaluer l'impact sur l'analyse, cf. la figure 5.6 où une observation et l'état vrai sont comparés. Tout le reste est identique à l'expérience précédente. Les erreurs de position sont simplement un décalage dans l'espace des observations, le décalage étant une variable aléatoire pour chaque observation, de moyenne nulle et de variance égale à 0,2.

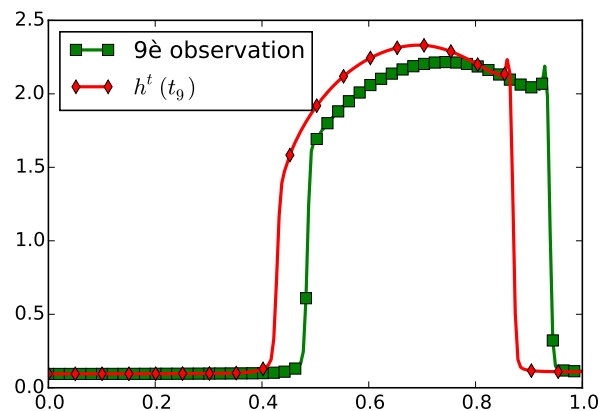
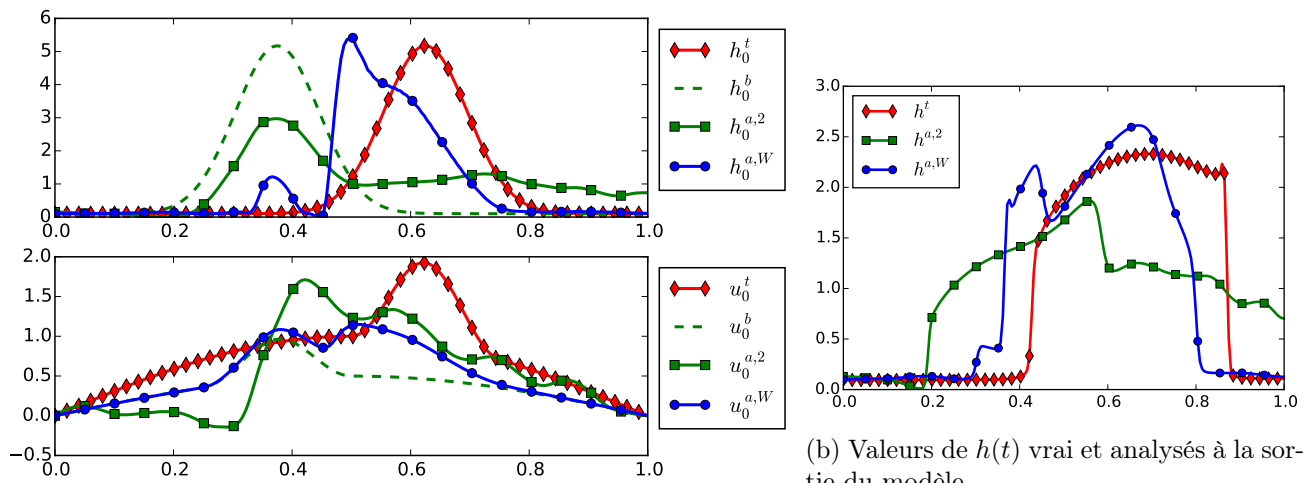


FIGURE 5.6 – Comparaison de la 9<sup>ème</sup> observation et de l'état vrai et temps  $t_9$  : l'observation contient des erreurs de position.

Comme dans le chapitre 4, nous nous attendons à ce que les erreurs de position n'influent pas vraiment sur la partie  $h_0$  de l'analyse de  $\mathcal{J}_W$ , contrairement à celle de l'analyse de  $\mathcal{J}_2$ . Concernant la vitesse  $u_0^{a,W}$ , celle-ci ne devrait pas être changée pour que  $h^{a,W}(t)$  soit proche en moyenne des observations.

Nous affichons les résultats dans la figure 5.7a. On constate en effet que l'analyse  $h_0^{a,W}$  n'a pas beaucoup bougé, mais que  $u_0^{a,W}$  est légèrement plus faible que dans l'expérience précédente sans erreur de position (cf. figure 5.5). Ceci est sans doute dû au fait que la première observation est en fait décalée à gauche de l'état vrai.

En tout cas, l'analyse pour  $\mathcal{J}_W$  a beaucoup moins changé que celle de  $\mathcal{J}_2$ , car, comme nous pouvons le voir,  $h_0^{a,2}$  est plus étalé. Ce résultat est positif pour  $\mathcal{J}_W$  dans le sens où les erreurs de position sur les observations n'ont pas beaucoup d'influence sur l'analyse, même dans le cas où une vitesse est contrôlée.



(a) En haut : Valeurs des différents  $h_0$  vrai, d'ébauche, et analysés.  
En bas : De même pour la vitesse  $u_0$ .

FIGURE 5.7 – Valeurs des différentes fonctions pour cette fois-ci l'assimilation sous erreurs de position sur les observations. Ici encore, seuls  $h_0$  et  $u_0$  sont contrôlés.

## 5.2 Contrôle du fond du bassin

Dans cette section nous contrôlerons uniquement  $f$  et  $h_0$ . La vitesse initiale  $u_0$  ne sera pas contrôlée ici, mais le sera dans la section suivante. Elle sera donc fixée égale à  $u_0^t$  qui sera posé comme étant nulle.

Le fond  $f$  est une fonction positive, de masse fixée à  $m_f$ . Ainsi, nous considérerons que  $f$  est un élément de  $m_f\mathcal{P}(\Omega)$ . On rappelle aussi que les observations sont des captures de l'altitude de la surface libre, donc des valeurs de  $h(t) + f$ . Étant donné qu'elles sont différentes de  $h$ , nous noterons les observations par la lettre  $y$ . De plus, ces observations sont des éléments de  $(1 + m_f)\mathcal{P}(\Omega)$ .

Nous utiliserons la distance de Wasserstein pour assimiler les observations. De plus, nous essaierons de retrouver des fonds  $f$  modélisés par des fonctions localisées, pour lesquelles la distance de Wasserstein semble adéquate à utiliser.

De même qu'avant, nous procéderons en deux étapes. Tout d'abord, seul  $f$  sera contrôlé, et l'on verra notamment quel produit scalaire convient le mieux. Ensuite, dans une deuxième étape, nous

contrôlerons  $f$  et  $h_0$  en utilisant aussi des termes d'ébauche.

### 5.2.1 Contrôle du fond seul, sans terme d'ébauche

Ici, nous cherchons à retrouver  $f$  uniquement et sans utiliser de terme d'ébauche. Pour ceci, nous utilisons la fonction coût de Wasserstein :

$$\mathcal{J}_W(f) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(f), y_i^{\text{obs}})^2$$

et dont nous comparerons les résultats avec ceux de la fonction coût  $\mathcal{L}^2$  :

$$\mathcal{J}_2(f) = \sum_{i=1}^{N^{\text{obs}}} \|\mathcal{G}_i(f) - y_i^{\text{obs}}\|^2.$$

Pour calculer le gradient de  $\mathcal{J}_W$  servant à minimiser  $\mathcal{J}_W$ , deux produits scalaires s'offrent à nous : le produit scalaire euclidien et celui de Wasserstein. Nous pouvons penser, suite aux résultats du chapitre 3, que la minimisation sera plus rapide avec le produit scalaire de Wasserstein. Nous étudions ceci dans cette section.

Les gradients de  $\mathcal{J}_W$  s'écrivent ainsi :

#### Proposition 5.2.1 (Gradient de $\mathcal{J}_W$ )

On a pour  $\eta_f$  une petite perturbation de  $f$  :

$$\mathcal{J}_W(f + \varepsilon \eta_f) = \mathcal{J}_W(f) + \varepsilon \left\langle \sum_i [\mathbf{M}_i^*[f] \cdot \Phi_i^{\text{obs}} + \Phi_i^{\text{obs}}], \eta_f \right\rangle_2 + \mathcal{O}(\varepsilon)$$

avec  $\Phi_i^{\text{obs}}$  le le potentiel de Kantorovich du transport entre  $\mathcal{G}_i(f)$  et  $y_i^{\text{obs}}$ , et  $\mathbf{M}_i^*[f]$  la composante en  $f$  du modèle adjoint, déterminé par les équations :

$$\mathbf{M}_i^* : \begin{cases} \partial_t \Phi + u \cdot \nabla \Phi + g \operatorname{div}(\Psi) = 0 \\ \partial_t \Psi + \operatorname{div}(\Psi)u + h \nabla \Phi = 0 \end{cases}$$

où  $\Phi(t_i) = \Phi_i^{\text{obs}}$  et  $\Psi(t_i) = 0$ . Alors,

$$\mathbf{M}_i^*[f] \cdot \Phi_i^{\text{obs}} := g \int_{[0, t_i]} \operatorname{div}(\Psi(t, x)) dt.$$

Donc, le gradient euclidien de  $\mathcal{J}$  s'écrit :

$$\operatorname{grad}_2 \mathcal{J}_W = \sum_i [\mathbf{M}_i^*[f] \cdot \Phi_i^{\text{obs}} + \Phi_i^{\text{obs}}]$$

et le gradient de Wasserstein s'écrit simplement :

$$\operatorname{grad}_W \mathcal{J}_W = -\operatorname{div}(f \nabla [\operatorname{grad}_2 \mathcal{J}_W]).$$

**Démonstration :** La démonstration est explicitée dans le théorème B.3.1. ■

Or, comme nous l'avons dit, le  $\Psi$  du modèle adjoint est homogène à  $-h\nabla\Phi$ . Ainsi,

$$\mathbf{M}_i^*[f].\Phi_i^{\text{obs}} \text{ est homogène à } -g \int_{[0,t_i]} \text{div}(h\nabla\Phi) dt$$

donc homogène à une densité, puisque  $\Phi$  est un potentiel de Kantorovich. Mais le gradient est lui la somme de  $\mathbf{M}_i^*[f].\Phi_i^{\text{obs}}$  et de  $\Phi_i^{\text{obs}}$ . Donc, la somme d'une densité et d'un potentiel de Kantorovich. Cela ne justifie le choix d'aucun des deux produits scalaires par rapport à l'autre (euclidien ou de Wasserstein).

Nous testons alors les deux gradients issus des deux produits scalaires. Nous les comparant en étudiant la minimisation de  $\mathcal{J}_{\mathcal{W}}$  avec chacun d'entre eux.

L'expérience consiste à retrouver  $f^t$  en minimisant  $\mathcal{J}_{\mathcal{W}}(f)$ . Deux des 20 observations ainsi que les  $h_0^t$  et  $f^t$  utilisés sont tracés sur la figure 5.8a. Nous arrêtons la minimisation après 40 itérations de la descente de plus forte pente.

On note par  $f^{a,W}$  l'analyse obtenue par  $\mathcal{J}_{\mathcal{W}}$  en utilisant le gradient de Wasserstein,  $f^{a,W^2}$  celle obtenue en utilisant le gradient euclidien, et  $f^{a,2}$  celle obtenue en utilisant la fonction coût  $\mathcal{J}_2$ . Ces analyses sont affichées sur la figure 5.8b dont les détails en  $x = 0.7$  sont affichés sur la figure 5.8c.

Nous pouvons observer sur cette figure que l'état vrai est bien retrouvé par  $\mathcal{J}_2$  et moins bien par  $\mathcal{J}_{\mathcal{W}}$ . Il y a notamment beaucoup plus d'oscillations dans le résultat obtenu à l'aide du gradient de Wasserstein, qui est dû au fait que ce gradient est une dérivée du gradient euclidien, puisque  $\text{grad}_{\mathcal{W}}\mathcal{J}_{\mathcal{W}} = -\text{div}(f\nabla\text{grad}_2\mathcal{J}_{\mathcal{W}})$ . Les dérivées ont en effet tendance à faire apparaître du détail.

De plus, nous comparons aussi les décroissances de  $\mathcal{W}_2(f^n, f^t)$  et de  $\|f^n - f^t\|_2$  en fonction de l'itération  $n$ , dans la figure 5.9. Nous constatons que le coût  $\mathcal{J}_2$  retrouve efficacement  $f^t$ . Par contre, pour le coût  $\mathcal{J}_{\mathcal{W}}$ , les deux gradients donnent des résultats équivalents en terme de métrique de Wasserstein. Nous ne fournissons pas d'analyse de ces résultats. C'est simplement une indication au fait que nous pourrions prendre les deux distances pour les termes d'écart à l'ébauche. En effet,

$$\mathcal{J}^o(f) + \|f - f^b\|^2$$

pourra être minimiser à l'aide d'un gradient euclidien qui minimisera efficacement le terme d'ébauche, et

$$\mathcal{J}^o(f) + \mathcal{W}_2(f, f^b)^2$$

avec un gradient de Wasserstein.

### 5.2.2 Contrôles du fond et de la hauteur de la colonne d'eau avec ébauche

Nous ajoutons maintenant le contrôle de la variable  $h_0$  en plus de celui de  $f$ . Par réalisme, nous mettons des termes d'écarts aux ébauches, ébauches possédant des erreurs de position. Pour les corriger, nous utiliserons la distance de Wasserstein. L'ébauche  $f^b$  est de masse  $m_f$ , tout comme  $f^t$ .

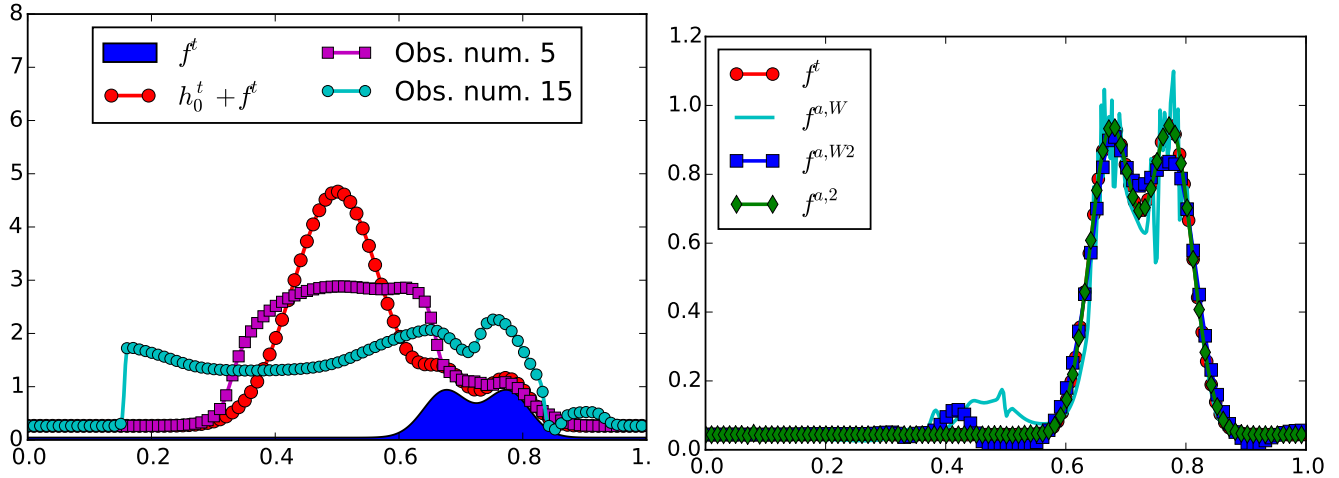
Aucune erreur sur les observations n'a été ajoutée.

Les  $h_0$  et  $f$  sont estimés en minimisant la fonction coût de Wasserstein suivante :

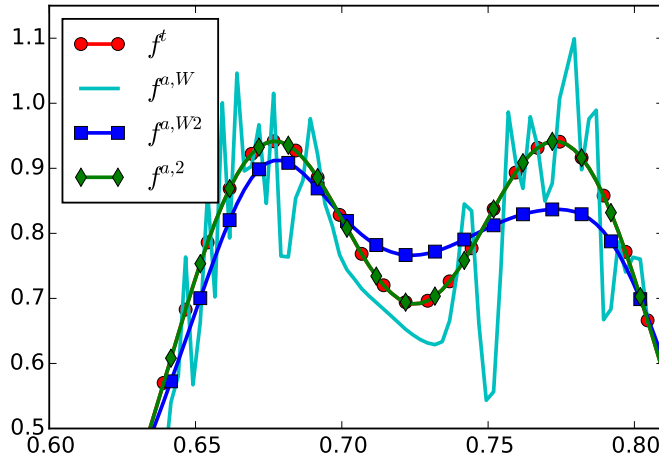
$$\mathcal{J}_{\mathcal{W}}(h_0, f) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(h_0, f), y_i^{\text{obs}})^2 + \omega_h \mathcal{W}_2(h_0, h_0^b)^2 + \omega_f \mathcal{W}_2(f, f^b)^2 + \mathcal{J}^r(h_0, f).$$

Nous avons ajouté des termes de régularisation  $\mathcal{H}^1$  pénalisant les fortes valeurs et variations de  $h_0$  et  $f$ . Nous nous attendons à ce que les analyses soient positionnées entre l'état vrai et l'ébauche.

Nous testons cette hypothèse dans l'exemple suivant. Nous avons pris des états vrais et des ébauches qui sont décalés, et pour  $h_0$ , et pour  $f$ , cf. la figure 5.10a. Dans la figure 5.10b on trace deux observations pour constater que l'obstacle posé au fond de l'eau gêne la progression de l'onde.



(a) Tracés de  $f^t$ ,  $h_0^t + f^t$  et des observations numéro 5 et 15 (il y a 20 observations en tout pour l'expérience). (b) Différentes analyses  $f^{a,W}$ ,  $f^{a,W2}$  et  $f^{a,2}$  obtenues.

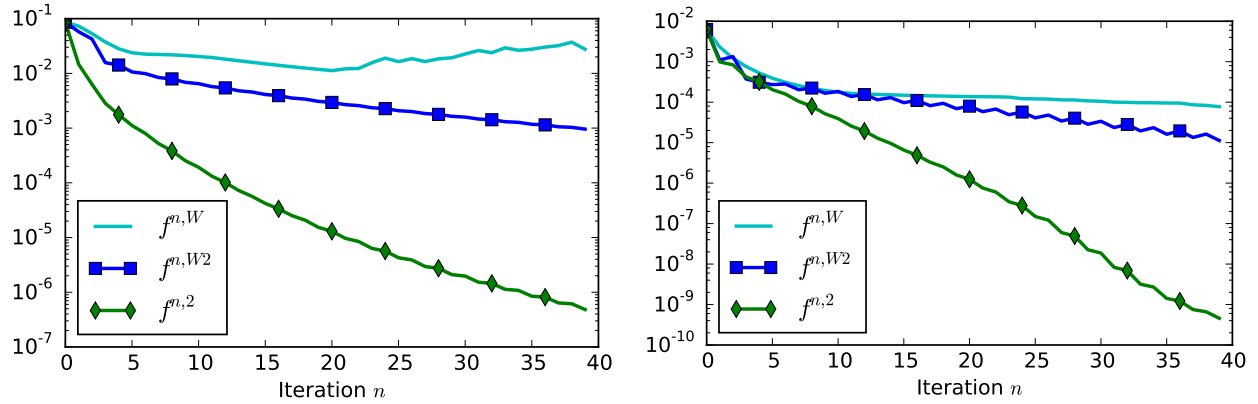


(c) Zoom de l'image (b) en  $x = 0.7$ .

FIGURE 5.8 – À gauche (a) : états vrais et quelques observations de la surface libre  $h + f$ , à droite (b) : états analysés en fonction des différents algorithmes, et (c) : zoom de la figure (b) en  $x = 0.7$ .

Étant donné que le terme d'écart à  $f^b$  dans  $\mathcal{J}_W$  est écrite avec une distance de Wasserstein, nous choisissons d'utiliser le gradient de Wasserstein pour minimiser  $\mathcal{J}_W$ . Après minimisation de  $\mathcal{J}_W$ , nous obtenons l'analyse tracée sur la figure 5.10a, dont les composantes sur  $h_0$  et  $f$  sont effectivement placées entre l'état vrai et l'ébauche. Nous pouvons comparer avec l'analyse issue de  $\mathcal{J}_2$ , également tracée sur la figure 5.10a, dont les composantes semblent écrasées à cause des termes d'ébauche. Pour les deux analyses, les écarts de position ou d'amplitude semblent bien rattrapés aux temps  $t$  supérieurs, cf. la figure 5.10c.

Le résultat de cette expérience est que même dans le cas où il y a plusieurs variables à assimiler, la propriété d'interpolation de Wasserstein semble respectée. Notamment, nous avons retrouvé un seul obstacle avec  $\mathcal{J}_W$ , alors que  $\mathcal{J}_2$  nous a donné un obstacle en deux parties. Cette bonne analyse est due au fait que les erreurs sur les données étaient de position.



(a) Valeur de  $\|f^n - f^t\|_2^2$  en fonction de la minimisation utilisée et de l'itération  $n$ . (b) Valeur de  $\mathcal{W}_2(f^t, f^n)^2$  en fonction de la minimisation utilisée et de l'itération  $n$ .

FIGURE 5.9 – Décroissances de la distance de  $f^n$  et  $f$ , en fonction de la distance (euclidienne à gauche, Wasserstein à droite), de l'itération  $n$  et de l'algorithme utilisé.

Nous avons mis un terme de Wasserstein pour l'ébauche sur  $f$  et vu que la minimisation s'est bien passée en utilisant un gradient Wasserstein pour  $f$ . Mais nous avons vu que le gradient euclidien était tout aussi capable de minimiser  $\mathcal{J}_W$ . C'est pourquoi nous nous disons que nous pouvons utiliser une distance  $\mathcal{L}^2$  pour calculer l'écart à  $f^b$  : c'est ce que nous voyons maintenant.

### 5.2.3 Utilisation d'une distance euclidienne pour le terme d'ébauche

Nous nous intéressons au cas où la fonction coût s'écrit :

$$\mathcal{J}_W(h_0, f) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}(h_0, f), y_i^{\text{obs}})^2 + \omega_h \mathcal{W}_2(h_0, h_0^b)^2 + \omega_f \|f - f^b\|_2^2.$$

Le terme d'écart  $f^b$  est écrit avec une distance euclidienne. Cette fonction pourrait être utilisée lorsque  $f^b$  est sujette à des erreurs de position plutôt que d'amplitude, et donc où l'on aurait besoin de mettre une distance  $\mathcal{L}^2$  sur  $f$ . On se demande si une telle fonction coût implique une contrainte  $\mathcal{L}^2$  sur  $f^b$ .

Nous avons vu dans la partie 5.2.1 que le produit scalaire  $\mathcal{L}^2$  sur  $f$  était compatible pour minimiser le terme d'écart aux observations. Ainsi, la minimisation de  $\mathcal{J}_W$  devrait bien se passer.

Nous testons les résultats de cette fonction coût en effectuant la même expérience que précédemment. La figure 5.11 montre l'analyse trouvée. Nous trouvons la même composante  $h_0^a$ . Par contre la composante du fond  $f^a$  a légèrement changé. Pour coller aux observations, on observe toujours une partie de la masse de  $f$  décalée vers l'état vrai, mais contrairement à avant, une partie de la masse de  $f$  essaie de coller à l'ébauche  $f^b$ .

Ici,  $f^a$  cherche à corriger à la fois la forme et la position de  $f^b$ , donnant une analyse de cette forme. Mettre un terme  $\mathcal{L}^2$  sur une des données est une contrainte sur la position de la solution, et nous constatons que ceci a tendance à dégrader sa qualité.

C'est un fait que nous n'avons pas relevé pour la vitesse  $u$  pour laquelle nous avons utilisé une distance euclidienne. Mais nous n'avons pas traité les erreurs de position sur la vitesse.

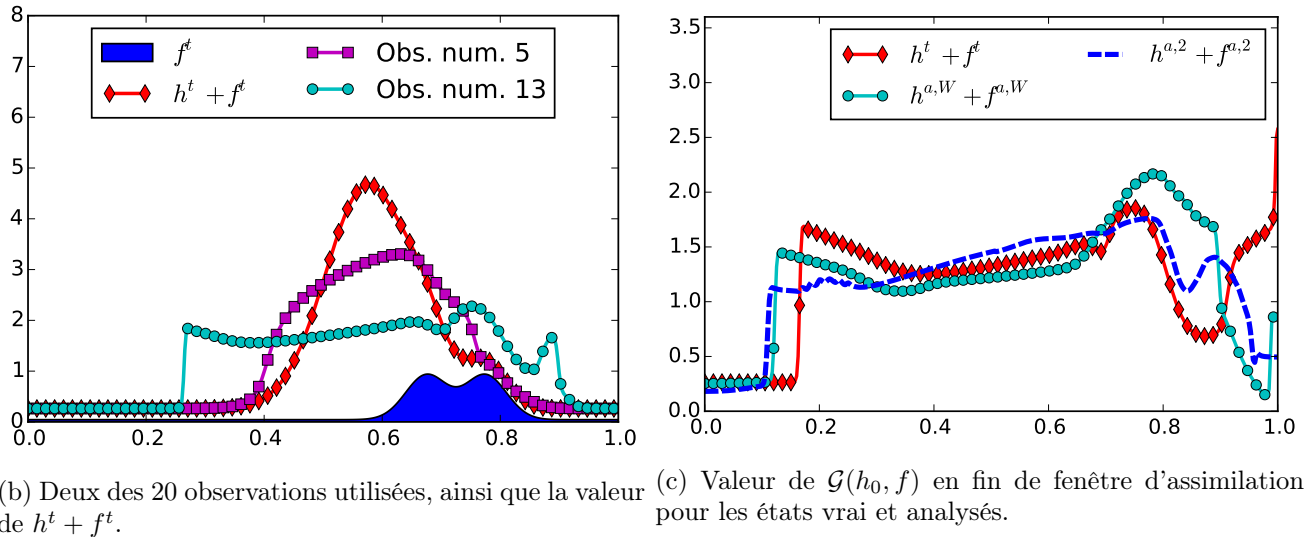
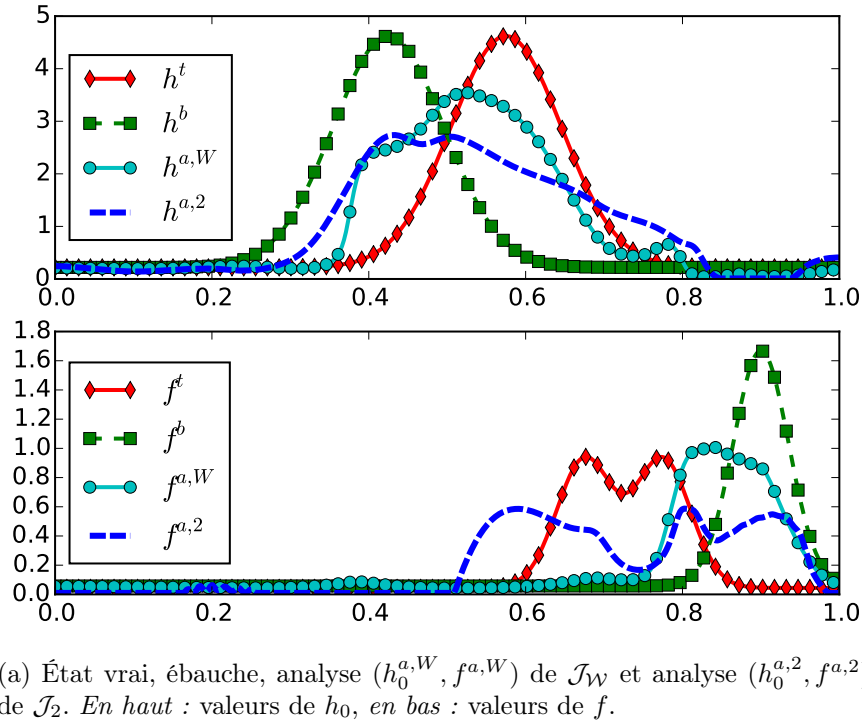


FIGURE 5.10 – Différents éléments impliqués dans l'expérience de minimisation de  $\mathcal{J}_W(h_0, f)$  et  $\mathcal{J}_2(h_0, f)$ . En (a) il y a les états vrai, l'ébauche et les deux analyses, en (b) quelques observations, et en (c) la comparaison des analyses et de la dernière observation.

### 5.3 Contrôle de toutes les variables en même temps

Nous effectuons dans cette section une expérience où nous contrôlons à la fois  $h_0$ ,  $u_0$  et  $f$ . Les états vrais et l'ébauche sont affichés sur la figure 5.12a. Nous avons mis  $u_0^b = 0$  pour mieux voir l'influence du terme d'ébauche sur la solution. Quelques observations sont également tracées dans la figure 5.12b.

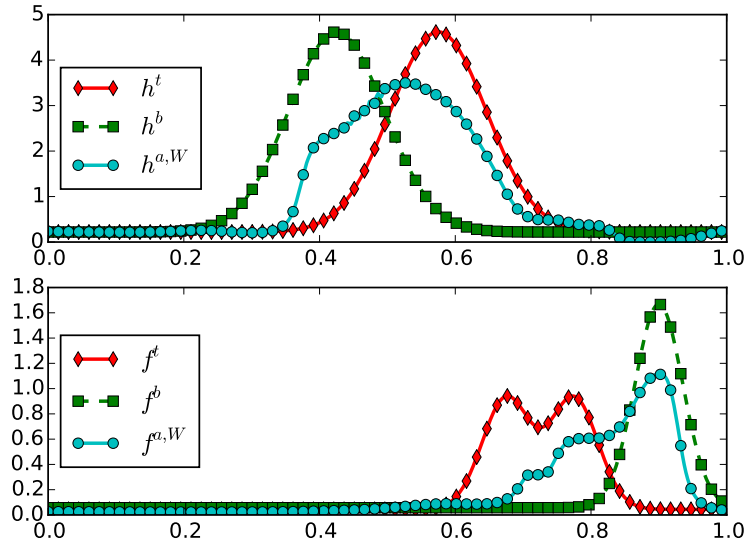
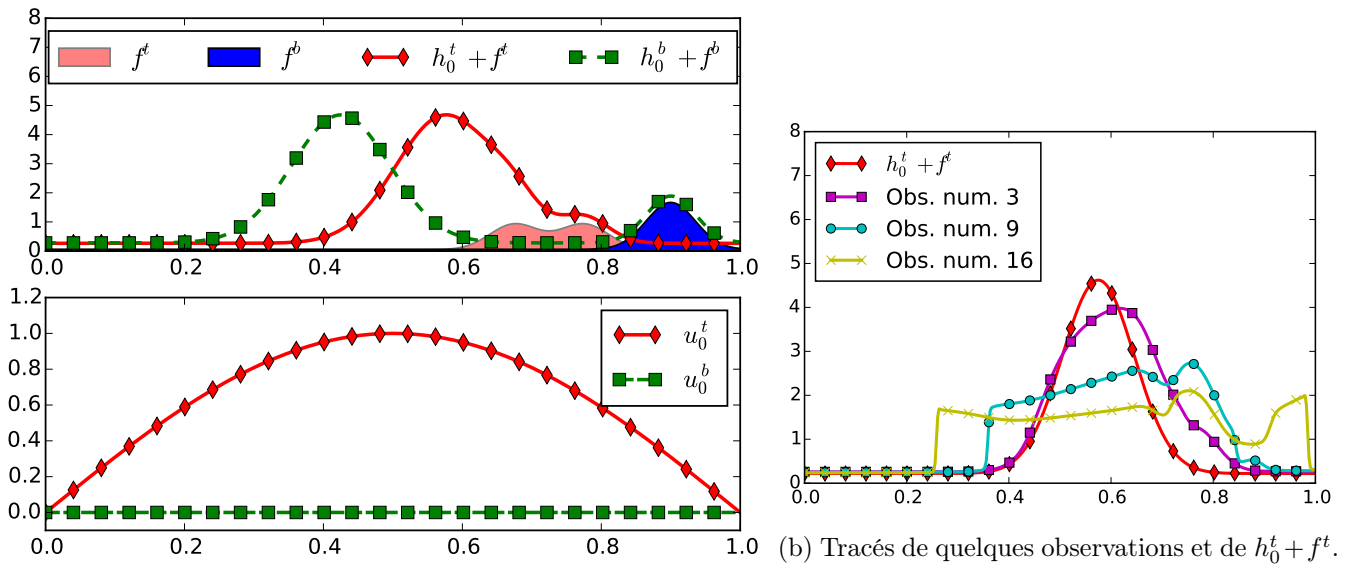

 FIGURE 5.11 – Analyse trouvée à l'issue de la minimisation de  $\mathcal{J}_W$ .

 (a) *En haut* : tracés de  $f$  (la hauteur du fond) et de  $h_0 + f$  (la hauteur de la surface libre) pour l'ébauche et l'état vrai d'où sont tirées les observations.

 FIGURE 5.12 – Éléments utilisés pour l'expérience, à gauche (a) :  $h_0$  et  $f$  en haut, et  $u_0$  en bas ; à droite (b) : trois des 20 observations.

Pour contrôler ces trois variables, la fonction coût à minimiser est :

$$\mathcal{J}_W(h_0, u_0, f) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(h_0, u_0, f), y_i^{\text{obs}})^2 + \omega_h \mathcal{W}_2(h_0, h_0^b)^2 + \omega_u \|u_0 - u_0^b\|^2 + \omega_f \mathcal{W}_2(f, f^b)^2 + \mathcal{J}^r(h_0, f).$$

Le terme de régularisation permet d'éviter la formation de fortes variations et de fortes valeurs dues à la tolérance de la distance de Wasserstein à ceux-ci. Concrètement, nous avons mis une



régularisation  $\mathcal{H}^1$  sur  $h_0$  et  $f$ .

Les observations sont des captures de  $h^t(t) + f^t$  aux temps  $t_i$ , sans introduction d'erreur.

Nous nous attendons avec une telle fonction coût à toujours avoir une correction en position des variables  $h_0$  et  $f$ . L'introduction du contrôle de la variable  $u_0$  va permettre plus de flexibilité sur la position de  $h_0$  et devrait être plus proche de l'ébauche (tout en étant plus près des observations grâce à la vitesse qui poussera le  $h_0$  initial sur les observations).

Les données ainsi que les analyses issues des minimisation de  $\mathcal{J}_W$  et de  $\mathcal{J}_2$  donnent les analyses que nous traçons dans la figure 5.13. De cette analyse nous pouvons reformuler les différentes conclusions obtenues avant. Nous observons en effet :

- une correction en position de la hauteur de la colonne d'eau initiale  $h_0$  ;
- une correction en amplitude de la vitesse, à l'endroit où est la colonne d'eau  $h_0^{a,W}$  ;
- une correction en position de l'obstacle au fond.

De plus, en comparant ces résultats avec ceux présentés dans la figure 5.10 où nous présentions les résultats d'une expérience similaire où la vitesse n'était pas contrôlée et était fixée égale à 0, nous constatons un léger décalage de la position de  $h_0^{a,W}$  vers l'ébauche.

La fonction coût  $\mathcal{J}_2$  retrouve elle  $h_0^{a,2}$  et  $u_0^{a,2}$  semblables à l'expérience de la figure 5.10, et de même que pour  $\mathcal{J}_W$ , une vitesse ayant tendance à pousser la masse initiale vers la droite. En tout cas, les vitesses initiales trouvées par  $\mathcal{J}_2$  et  $\mathcal{J}_W$  sont de même amplitude.

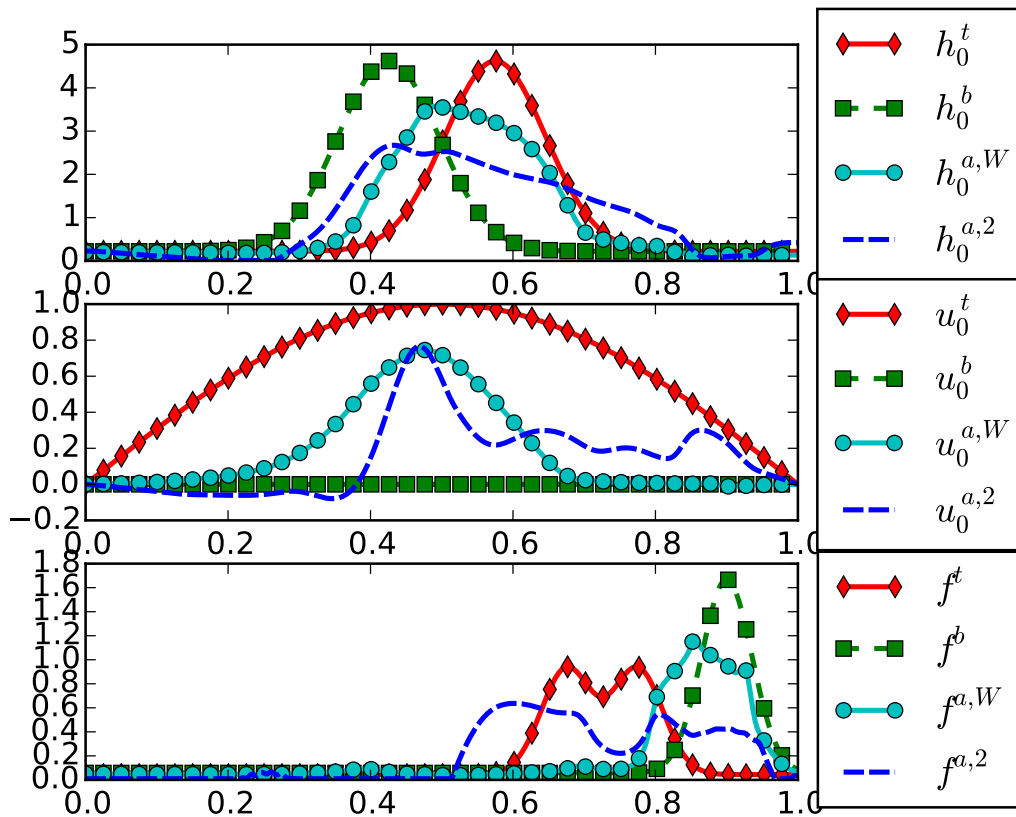


FIGURE 5.13 – Analyse trouvée par la minimisation de  $\mathcal{J}_W$ , en haut :  $h_0$ , au milieu :  $u_0$  et en bas :  $f$ .

Nous avons réussi à minimiser  $\mathcal{J}_{\mathcal{W}}(h_0, u_0, f)$  et à faire ressortir les différentes conclusions que l'on recense dans la section suivante.

## 5.4 Conclusion générale de ces expériences

Nous avons effectué une assimilation de données dans un modèle simple de Saint-Venant avec fond, où trois variables interviennent : la hauteur de la colonne d'eau  $h(t)$  de valeur initiale  $h_0$ , la vitesse  $u(t)$  de valeur initiale  $u_0$  et le fond  $f$ . Nous avons supposé que  $h_0 \in \mathcal{P}(\Omega)$  et que  $f \in m_f \mathcal{P}(\Omega)$  pour pouvoir utiliser une distance de Wasserstein. Notons que nous aurions pu simplement poser  $h_0 \geq 0$  et  $f \geq 0$  et utiliser une distance  $WF$ , chose que nous n'avons pas faite par simplicité et par conviction que les résultats issus d'une distance de Wasserstein ou d'une distance  $WF$  sont équivalents. De plus, nous avons choisi la distance de Wasserstein car elle est beaucoup plus rapide à calculer en 1D.

Les résultats issus d'une fonction coût  $\mathcal{J}_{\mathcal{W}}$  ont montré plusieurs caractéristiques de l'assimilation de données utilisant le transport optimal.

- Les résultats sur les variables de  $\mathcal{P}(\Omega)$  contrôlées à l'aide de la distance de Wasserstein (dans l'ébauche notamment) sont équivalents à ceux trouvés dans le chapitre 4 : c'est-à-dire que par rapport à l'ébauche, l'analyse est une correction de la position. Par contre, les détails ne sont pas retrouvés, et ce parce que nous avons notamment mis des termes de régularisation.
- La vitesse, variable n'appartenant pas à  $\mathcal{P}(\Omega)$ , a également été retrouvée. Nous avons vu notamment à travers l'analyse du gradient que les erreurs de positions des variables observées étaient corrigées par des corrections d'amplitude sur la vitesse. Nous n'avons par contre pas traité le cas où la vitesse avait des erreurs de position : par exemple si l'ébauche a des grandes vitesses à un endroit mal positionné par rapport à la réalité, nous n'avons pas pu faire une correction de la position de cette structure dans l'ébauche, à l'aide d'une distance de Wasserstein par exemple.
- Nous avons vu, en analysant le gradient euclidien, quel produit scalaire choisir pour écrire le gradient servant à minimiser une fonction coût  $\mathcal{J}_{\mathcal{W}}$  écrite sans terme d'écart aux ébauches. Ce produit scalaire nous guide pour le choix de la distance à écrire pour l'ébauche : en effet, nous avons vu qu'il pouvait être difficile de minimiser avec un tel produit scalaire un terme écrit avec une autre distance.
- Pour comparer avec les résultats issus de  $\mathcal{J}_2$ , nous pouvons dire que les résultats des variables traitées avec la distance de Wasserstein ( $h_0$  et  $f$ ) sont de même forme que l'ébauche mal placée, contrairement aux composantes de l'analyse issue de  $\mathcal{J}_2$  qui semblent elles écrasées, sur un support plus large. Les résultats concernant la vitesse sont les mêmes, c'est-à-dire un entre-deux entre l'ébauche et ce qu'il faut pour corriger parfaitement les écarts entre l'analyse et les observations.
- En calculant un écart à l'ébauche  $f^b$  avec la distance euclidienne plutôt que la distance de Wasserstein, nous avons vu une correction double : une correction de l'amplitude (par rapport à  $f^b$ ) en plus d'une correction de la position (pour les observations), ce qui crée une analyse en deux parties et donc de forme peu souhaitable.

Ceci conclut notre analyse sur  $\mathcal{J}_{\mathcal{W}}$ . Ces expériences, loin d'être exhaustives, nous ont permis d'accumuler un certain nombre de résultats et de conclusions sur l'utilisation de la distance de Wasserstein pour l'assimilation de données. Nous concluons ainsi le travail réalisé en résumant les résultats acquis et en l'ouvrant à des possibilités d'applications concrètes dans le chapitre suivant.



# Chapitre 6

## Conclusion et perspectives

### Contenu du chapitre

6.1	Résumé du travail effectué . . . . .	127
6.2	Perspectives . . . . .	131
6.2.1	Produit scalaire alternatif pour le fond . . . . .	131
6.2.2	$n$ -dimension . . . . .	132
6.2.3	Statistiques d'erreurs . . . . .	135
6.2.4	Occultations . . . . .	138
6.2.5	Assimilation de données par filtrage et par la distance de Wasserstein . . . . .	139

### 6.1 Résumé du travail effectué

L'assimilation variationnelle de données consiste à estimer l'état d'un système physique en minimisant une fonction coût. Cette fonction coût calcule la somme d'écart entre l'état analysé et les différentes données, écarts définis en choisissant convenablement une distance. Nous nous sommes placés dans le cas où les observations sont des images ou des données denses en général, c'est-à-dire des données structurées en espace, qu'on supposera positifs ou éléments de  $\mathcal{P}(\Omega)$ , et avec lesquelles des erreurs de position pouvaient intervenir. Nous avons vu à travers plusieurs expériences que choisir une distance euclidienne pouvait impliquer des analyses définies sur un support trop large. Ceci a des conséquences notamment pour la prévision du système, car les effets sont lissés (corrections d'amplitude réduite), ou au contraire amplifiés (vitesse largement surestimée en certains endroits pour compenser cette amplitude réduite).

Pour pouvoir mieux gérer les erreurs de position, nous avons choisi d'utiliser une distance de Wasserstein, car cette distance pourrait convenablement gérer les erreurs de position, notamment en fournissant un barycentre de position moyenne aux données. Nous écrivons alors l'écart aux observations comme :

$$\mathcal{J}^o(\mathbf{x}_0) = \sum_{i=1}^{N^{\text{obs}}} \mathcal{W}_2(\mathcal{G}_i(\mathbf{x}_0), y_i^{\text{obs}})^2.$$

Ici,  $\mathcal{G}_i$  transforme  $\mathbf{x}_0$  en son équivalent-observation numéro  $i$ . Nous supposons que  $\mathcal{G}_i(\mathbf{x}_0)$ ,  $y_i^{\text{obs}}$  sont des éléments de  $\mathcal{P}(\Omega)$  pour pouvoir utiliser  $\mathcal{W}_2$ . Nous pouvons relâcher la contrainte sur la valeur de l'intégrale des données et utiliser la distance  $WF$ .

En assimilation variationnelle de données il est classique d'ajouter des régularisations dites de Tikhonov qui sont des termes d'écart à l'ébauche, puisque les observations seules ne suffisent pas à retrouver  $\mathbf{x}_0$  (problème sous-déterminé), et puisque les ébauches comportent des informations utiles à incorporer dans l'analyse. Nous écrivons de manière générale le terme d'écart à l'ébauche :

$$\mathcal{J}^b(\mathbf{x}_0) = \sum_{j=1}^{N_{\text{var}}} d_j(\mathbf{x}_0^j, \mathbf{x}_0^{b,j})^2$$

avec  $\mathbf{x}_0^j$  et  $\mathbf{x}_0^{b,j}$  les  $j$ -èmes variables physiques de  $\mathbf{x}_0$  et  $\mathbf{x}_0^b$ , et  $d_j$  la distance à choisir pour cette variable. Nous avons vu qu'il arrivait que les ébauches  $\mathbf{x}_0^{b,j}$  peuvent être entachées d'erreurs de position notamment parce qu'elles sont issues d'un calcul précédent. Pour les corriger, nous avons choisi d'utiliser des distances de Wasserstein, mais seulement sur les variables de  $\mathcal{P}(\Omega)$ . Si ces variables sont simplement positives, nous pouvons utiliser une distance  $WF$  dont les résultats sont équivalents à ceux fournis par  $\mathcal{W}_2$ . Nous utilisons des distances  $\mathcal{L}^2$  sur les autres types de variables.

La fonction coût à minimiser est alors la somme des deux écarts :

$$\mathcal{J}_{\mathcal{W}}(\mathbf{x}_0) = \mathcal{J}^o(\mathbf{x}_0) + \mathcal{J}^b(\mathbf{x}_0).$$

Minimiser de telles fonctions coûts revient à vouloir trouver une analyse qui soit à la fois :

- proche en position aux observations ;
- proche en position aux composantes de l'ébauche avec lesquelles l'écart est mesuré à l'aide de la distance de Wasserstein ;
- et proche en amplitude aux composantes de l'ébauche dont l'écart est mesuré à l'aide de la distance euclidienne.

Nous allons parler de cette fonction  $\mathcal{J}_{\mathcal{W}}$ , mais avant traitons de la fonction  $\mathcal{J}_{\mathcal{H}}$  qui se définit comme  $\mathcal{J}_{\mathcal{W}}$  mais avec des distances  $\mathcal{H}^{-1}$  à la place des distances de Wasserstein. Nous avons voulu écrire cette fonction coût puisque la distance  $\mathcal{H}^{-1}$  tient compte de l'écart entre deux objets. Cependant, les propriétés d'interpolation sont les mêmes que pour la distance euclidienne si bien que sous erreur de position, nous nous sommes retrouvés avec des analyses de support plus large que ceux des données. Notre principal problème a été la minimisation de  $\mathcal{J}_{\mathcal{H}}$  puisqu'elle nécessite une projection  $\mathcal{H}^{-1}$  sur  $\mathcal{P}(\Omega)$ , projection que nous n'avons pas pu définir.

Quelques expériences sur un modèle Saint-Venant 1D ont été menées pour étudier l'impact de  $\mathcal{J}_{\mathcal{W}}$ . Ce modèle simple a pour intérêt d'être non-linéaire et d'offrir des transports de masses. Il fait intervenir la hauteur de la colonne d'eau  $h$  et la vitesse  $u$ , de valeurs initiales respectives  $h_0$  et  $u_0$ . Nous avons agrémenté le modèle d'une variable  $f$  modélisant le fond et par exemple des obstacles qui gênaient la propagation des ondes.

Pour avoir un impact sur la distance de Wasserstein, nous avons choisi des  $h_0$  à support compact et restreint. L'effet dans le modèle est d'avoir des ondes dispersives et s'atténuant très vite. Nous n'avons donc pas réalisé d'expériences en temps long, mais nous avons tout de même pu établir quelques résultats.

Avant de les résumer, rappelons la minimisation numérique de  $\mathcal{J}_{\mathcal{W}}$ . Ces minimisations ont été réalisées avec une descente de plus forte pente, qui est un algorithme de minimisation itératif de la forme  $\mathbf{x}_0^{n+1} = \mathbf{x}_0^n + \alpha^n d^n$ , avec  $\alpha^n$  un scalaire positif et  $d^n$  une direction de descente. La descente

de plus forte pente consiste à prendre pour  $d^n$  l'opposée d'un gradient de  $\mathcal{J}_W$ . Il est également nécessaire de projeter les variables concernées sur  $\mathcal{P}(\Omega)$ .

Pour cette méthode il est nécessaire de calculer le gradient. Le calcul du gradient euclidien de  $\mathcal{J}_W$  repose sur la différentiation de la distance de Wasserstein donnée par le potentiel de Kantorovich, et sur la méthode de l'adjoint. Ayant ce gradient euclidien, nous avons pu effectuer la descente. Il a parfois été judicieux de choisir le gradient de Wasserstein à la place (défini à l'aide du produit scalaire de Wasserstein) pour la descente, dans les expériences du chapitre 4 par exemple. Utiliser ce gradient permet d'avoir une convergence beaucoup plus rapide. De plus, nous avons écrit la descente pour une variable  $\rho_0$  par :

$$\rho_0^{n+1} = (I - \alpha^n \nabla \Phi^n) \# \rho_0^n$$

avec  $\alpha^n \geq 0$ ,  $\Phi^n$  le gradient euclidien de  $\mathcal{J}_W$ . Pour le calcul de  $\rho_0^{n+1}$  en fonction de  $\rho_0^n$ , nous avons également mis un terme dissipatif pondéré par  $\nu^n$ , ce qui a pour effet de ne pas tomber dans des minima locaux. Nous faisons décroître la valeur de  $\nu^n$  en fonction des itérations.

Les résultats décrits ci-dessus peuvent être étendus à des algorithmes de minimisation plus rapides tels que la méthode du gradient conjugué. Il faudrait alors les écrire en utilisant le produit scalaire de Wasserstein. Nous avons réussi à le faire avec le produit scalaire  $\mathcal{H}^{-1}$  pour minimiser des fonctions écrites avec la distance  $\mathcal{H}^{-1}$ , ce qui donne des convergences rapides.

Enfin, pour étudier l'impact des différentes variables dans l'écart aux observations  $\mathcal{J}^o$ , nous avons choisi d'étudier la nature du gradient de  $\mathcal{J}^o$ . Il nous a été salutaire d'étudier ce gradient pour mettre en évidence une incohérence dans les données. En effet, nous ajoutions des potentiels de Kantorovich à des densités dans une minimisation utilisant le gradient euclidien (cf. le chapitre 3), ce qui freine la convergence. À la place, nous avons pris un gradient de Wasserstein pour minimiser plus efficacement. Nous avons fait de même avec la vitesse pour trouver qu'un gradient euclidien convenait bien, puisque dans le cas où il n'y a pas de terme d'ébauche et que le minimum des fonctions coûts est atteint en  $u_0^t$ , nous avions une meilleure convergence en termes de  $\|u_0^n - u_0^t\|_2$  pour  $\mathcal{J}_W$  que pour  $\mathcal{J}_2$ .

Cette étude nous a permis aussi d'identifier les bonnes distances pour calculer le terme d'écart à l'ébauche. Nous avons en effet pointé des inconsistances en utilisant une distance sur l'ébauche différente de la distance aux observations dans le cas du 3D-Var, cf. la section 4.1. En effet, l'équation que vérifie le minimiseur est l'équation d'Euler-Lagrange et est la nullité du gradient :

$$\text{grad}\mathcal{J}(\mathbf{x}_0) = \text{grad}\mathcal{J}^o(\mathbf{x}_0) + \text{grad}\mathcal{J}^b(\mathbf{x}_0) = 0$$

et nous avons montré des exemples de minimiseurs peu satisfaisants parce que les deux parties du gradient,  $\text{grad}\mathcal{J}^o(\mathbf{x}_0)$  et  $\text{grad}\mathcal{J}^b(\mathbf{x}_0)$ , n'étaient pas de même nature. Ainsi nous avons choisi de prendre un  $\mathcal{J}^b$  qui ne soit pas inconsistant avec  $\mathcal{J}^o$ .

Les expériences que nous avons menées dans les chapitres 4 et 5 sont assez simplistes : modèle 1D, nombreuses observations, totales et souvent parfaites, fenêtre d'assimilation courte, pas de matrice de covariance d'erreurs. Elles nous ont cependant conduit à différentes analyses, que nous donnons ici en toute généralité.

- Globalement, les erreurs de position dans les observations et dans les variables de l'ébauche avec lesquelles l'écart est mesuré par  $\mathcal{W}_2$  sont moyennées : c'est-à-dire que nous retrouvons une analyse de position moyenne à celle de toutes les données et d'amplitude équivalente. C'est ce que nous cherchions à retrouver au début de ce travail et l'approche que nous proposons répond bien à la problématique initiale.

Si nous comparons avec l'analyse trouvée par  $\mathcal{J}_2$ , sous erreurs de position cette dernière se retrouve en fait beaucoup avec un support beaucoup plus large que celui des données. L'analyse de  $\mathcal{J}_W$  est donc d'amplitude et de largeur plus satisfaisante que celle de  $\mathcal{J}_2$ . L'intérêt se retrouve par exemple dans une assimilation dans un modèle d'évolution d'un cyclone, qui, s'il y a des erreurs de position dans les données, va fournir plusieurs cyclones de faible amplitude pour  $\mathcal{J}_2$  ou un seul cyclone pour  $\mathcal{J}_W$ .

Nous pouvons cependant formuler quelques critiques associées à ces résultats :

- La forme de l'analyse n'est généralement pas retrouvée, même si les données sont toutes de mêmes formes. Nous retrouvons une analyse sans détails particuliers. En effet, sauf si ces détails ont un impact significatif dans l'analyse à  $t > 0$ , ceux-ci ne sont pas significatifs pour la distance de Wasserstein.
  - Les formes prises par l'analyse de  $\mathcal{J}_W$  sont souvent très oscillantes ou présentent des Diracs, ce qui n'est pas souhaitable. Des termes de régularisation peuvent être ajoutés pour pallier ce problème, amplifiant encore l'effet lissant de  $\mathcal{J}_W$ .
  - Le fait d'avoir des masses localisées dans les différentes données contribue aux bons résultats donnés par la distance de Wasserstein. En effet, les barycentres de telles masses sont encore des masses localisées. En revanche, comme illustré par la figure 2.10, le barycentre de deux détails mal placés sur une forme ne bougeant pas ressemble au barycentre euclidien. Si nous devons assimiler des données de température par exemple, variant de  $\sim 250 K$  à  $\sim 300 K$ , sujettes à des erreurs de position, les structures présentes ne seraient pas interpolées de la bonne façon par la distance de Wasserstein. Une façon de contourner le problème est de soustraire une fonction commune à toutes les données pour que les structures ressortent.
  - Lorsque nous utilisons des distances euclidiennes pour calculer l'écart à l'ébauche, par choix ou faute de pouvoir en utiliser d'autres comme pour la vitesse, cela crée généralement des analyses aux formes bizarres, avec une correction de la position combinée à une correction de l'amplitude de l'ébauche. On réfère le lecteur à la figure 4.1d notamment où les deux effets apparaissent. Cela peut créer des effets indésirables qu'il faudra prendre compte, par exemple en baissant le poids sur de tels termes d'ébauche.
- Nous n'avons pas trouvé de distances gérant l'écart de position pour des variables non positives, comme la vitesse. Un changement de variable pourra être tout de même envisagé avant d'utiliser la distance de Wasserstein. Nous pensons qu'il est possible de faire des changements de variable non bijectifs, l'idée est juste de faire ressortir la position de l'information à interpoler.

- Nous avons également trouvé des résultats de convergence intéressants dans le cas de contrôle de la variable vitesse seule à partir d'observations de la hauteur d'eau, à l'aide de la distance de Wasserstein. Nous avons en effet obtenu une convergence plus rapide en utilisant la distance de Wasserstein qu'en utilisant une distance euclidienne. Cette vitesse de convergence est à nuancer avec le temps de calcul de la distance de Wasserstein en dimension supérieure notamment, mais elle indique que la fonction coût  $\mathcal{J}_W(u_0)$  est probablement plus convexe que  $\mathcal{J}_2(u_0)$ .

Le transport optimal peut ainsi en théorie trouver des applications dans le cadre de l'assimilation de données entachées d'erreurs de position. Cependant, pour plus de réalisme, il faudra réaliser un travail supplémentaire d'ouverture dont nous présentons les grandes lignes à présent.

## 6.2 Perspectives

### 6.2.1 Produit scalaire alternatif pour le fond

En voulant assimiler le fond  $f$  du bassin, dans la section 5.2, nous avons vu que le gradient (euclidien) était la somme d'un potentiel de Kantorovich et d'une densité, qu'on écrit comme

$$\text{grad}_2 \mathcal{J}(f) = -\text{div}(h\nabla\Phi) + \Phi'$$

avec  $\Phi$  et  $\Phi'$  des fonctions de même nature qu'un potentiel de Kantorovich. Le gradient est ainsi le mélange d'une partie  $\mathcal{L}^2$  et de Wasserstein. Un changement d'amplitude sur  $f$  implique donc un changement à la fois d'amplitude et de position sur les observations. Physiquement, nous interprétons ceci par le fait qu'un changement d'amplitude du fond  $f$  joue sur la taille des obstacles, pouvant potentiellement freiner ou libérer la circulation de l'eau, créant donc des déplacements relatifs de la masse.

Le gradient euclidien est la somme d'une densité et d'un potentiel de Kantorovich, rendant le choix du produit scalaire difficile : le gradient doit-il être défini par le produit scalaire euclidien ou de Wasserstein ? Le gradient de Wasserstein est défini comme :

$$\text{grad}_W \mathcal{J} = -\text{div}(f\nabla\text{grad}_2 \mathcal{J}).$$

Pour circonvier ce problème, on peut utiliser le produit scalaire intermédiaire aux deux qui permet de récupérer le potentiel. Le produit scalaire intermédiaire à  $\mathcal{L}^2$  et  $\mathcal{W}_2$ , défini en toute généralité pour  $\rho_0 \in \mathcal{P}(\Omega)$  et pour  $\eta, \eta' \in T_{\rho_0} \mathcal{P}$  tels que :

$$\eta = -\text{div}(\rho_0 \nabla \Phi), \quad \eta' = -\text{div}(\rho_0 \nabla \Phi') \text{ avec conditions aux bords de Neumann,}$$

par :

$$\begin{aligned} \langle \eta, \eta' \rangle_\tau &= \int_\Omega \rho_0 \nabla \Phi \cdot \nabla \Phi' \, dx + \tau \int_\Omega \eta \eta' \, dx \\ &= \int_\Omega [\Phi' + \tau \eta'] \eta \, dx \end{aligned}$$

Pour  $\tau = 0$  nous retrouvons le produit scalaire de Wasserstein et pour  $\tau \rightarrow \infty$  un équivalent au produit scalaire euclidien. Pour ce produit scalaire, le gradient est donné en fonction du gradient euclidien par :

$$\text{grad}_\tau \mathcal{J} = -\text{div}(\rho_0 \nabla \Phi) \text{ avec } \Phi \text{ tel que } \Phi - \tau \text{div}(\rho_0 \nabla \Phi) = \text{grad}_2 \mathcal{J}. \quad (6.2.1)$$

En effet, avec ce gradient-ci, nous avons

$$\begin{aligned} \langle \text{grad}_\tau \mathcal{J}, \eta \rangle_\tau &= \int_\Omega [\Phi - \tau \text{div}(\rho_0 \nabla \Phi)] \eta \, dx \\ &= \langle \Phi - \tau \text{div}(\rho_0 \nabla \Phi), \eta \rangle_2 \end{aligned}$$

et par identification avec la différentielle de  $\mathcal{J}$  qui s'écrit aussi  $\langle \text{grad}_2 \mathcal{J}, \eta \rangle_2$ , nous prouvons cette valeur du gradient. Ainsi, lorsque nous regardons la définition (6.2.1) du gradient, nous récupérerons ainsi un potentiel  $\Phi$  d'une forme étant la somme d'un potentiel de Kantorovich et d'une densité, comme le gradient euclidien de  $\mathcal{J}_W(f)$ . Le gradient issu de ce produit scalaire, pour un certain  $\tau$ , pourrait donner une meilleure convergence.



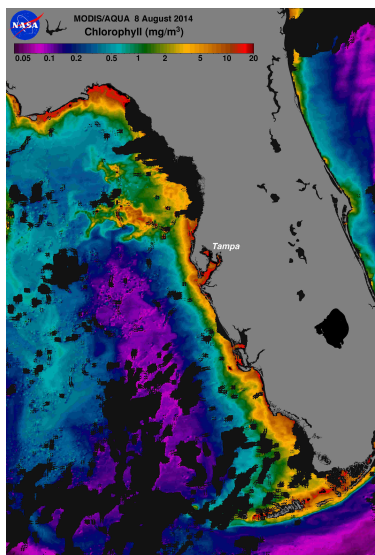
Ce paragraphe dit qu'en théorie nous pouvons mélanger les produits scalaires ce qui ouvre la voie à beaucoup de possibilités de minimisation, basées sur l'étude du gradient.

### 6.2.2 $n$ -dimension

Un des facteurs limitants de notre étude est la contrainte d'être en 1 dimension. Nous n'avons pas effectué de tests en dimension supérieure par simplicité. Cependant, nous affirmons que les résultats obtenus sont généralisables en dimension supérieure.

En 2 et 3 dimensions, le transport optimal pour l'assimilation de données trouverait plus de modèles intéressants et d'applications. Nous pouvons citer quelques modèles où les positions peuvent jouer.

En océanographie, différents traceurs sont utilisés pour contrôler l'état de l'océan. Ces traceurs sont advectés dans les modèles et leurs déplacements donnent une information sur la dynamique du système. Nous pouvons noter la concentration de chlorophylle provenant des phytoplanctons (cf. figure 6.1a) ou la concentration des polluants (cf. figure 6.1b). En météorologie, des modèles de prévision d'un cyclone (cf. une représentation idéalisée dans la figure 6.2) sont appropriés pour la distance de Wasserstein puisqu'il y a transport d'une masse localisée (le cyclone). En imagerie médicale, retrouver la position des organes dans le corps humain à partir des données du scanner pourrait faire appel à la distance de Wasserstein.



(a) Concentration de chlorophylle dans l'océan autour de la Floride. Les tâches noires correspondent à la présence de nuage.



(b) Photo d'avion d'une nappe de pétrole dans l'océan. Cette nappe peut servir à contrôler l'état de l'océan, mais on peut aussi contrôler sa position initiale.

FIGURE 6.1 – Exemples de traceurs utilisés en océanographie.

Un autre point à étudier sont les algorithmes de calcul du transport optimal. Nous avons utilisé des algorithmes de transport optimal en 1D dont le calcul s'effectue en quelques centièmes de seconde et dont le résultat était exact. En dimension supérieure, cet algorithme n'est pas utilisable. Les algorithmes alternatifs peuvent présenter plusieurs défauts, en particulier ils ne sont qu'une approximation et ils sont longs à calculer.

À notre connaissance seuls deux algorithmes sont utilisables en dimensions supérieures. Nous pourrions citer aussi les algorithmes type algorithme Hongrois ou issu de [Mérigot, 2011] mais qui

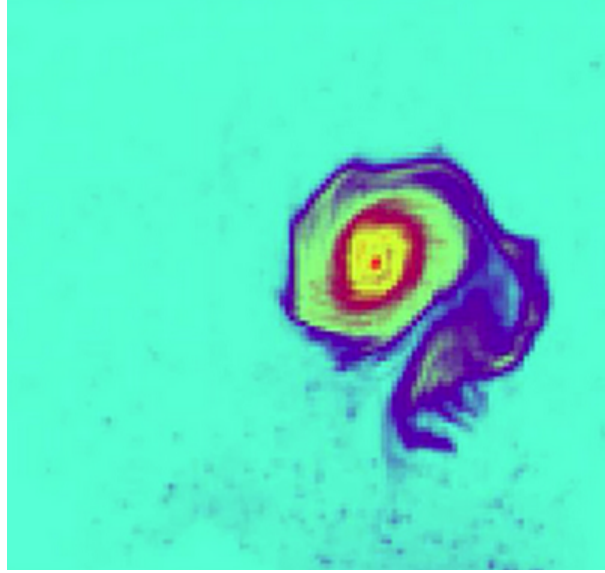


FIGURE 6.2 – Image d’un tourbillon issue de la plate-forme tournante Coriolis de Grenoble. Source : [Vidard et al., 2014].

ne semblent pas pouvoir calculer  $WF$ . Ces deux algorithmes sont :

- L’algorithme primal-dual, cf. la section A.5 de l’annexe. Cet algorithme a la propriété de calculer exactement le transport optimal à convergence. Il crée par itérations une suite de variables  $(U^n, V^n, W^n)$  qui convergent vers les solutions des problèmes primaux et duaux. À chaque itération, un laplacien discrétisé sur une grille spatio-temporelle doit être résolu et c’est ce qui prend le plus de temps. De plus, étant donné qu’on ne va pas à convergence, il faut étudier la proximité de la solution donnée à la vraie solution. Nous faisons ceci ici.
- L’algorithme de régularisation entropique, donné dans [Cuturi, 2013] pour  $\mathcal{W}_2$  et présenté dans l’annexe A.6, et donné dans [Chizat et al., 2016] pour  $WF$  (non implémenté ici). Cet algorithme calcule une approximation de la distance de Wasserstein, et est très rapide mais dépend fortement d’un paramètre  $\varepsilon > 0$  dont le réglage est délicat. Nous l’étudions ici pour son utilisation en dimension supérieure.

Nous effectuons une expérience dans laquelle ces deux autres algorithmes sont testés. Nous effectuons un test simple de minimisation de la fonction coût

$$\mathcal{J}_{\mathcal{W}}(h_0) = \sum_i \mathcal{W}_2(\mathcal{M}_i(h_0), h_i^{\text{obs}})^2 + \mathcal{W}_2(h_0, h_0^b)^2$$

avec  $\mathcal{W}_2$  calculé numériquement avec les différents algorithmes. Le modèle utilisé est le modèle de Saint-Venant, ce test simple est similaire aux expériences menées dans le chapitre 4.

Nous étudions la qualité des différents résultats. Nous affichons sur la figure 6.3 les différentes analyses obtenues et dans la table 6.1 les performances des différents algorithmes, niveau temps total de l’assimilation, qualité du résultat en fonction de  $\mathcal{J}$  utilisant l’algorithme exact, et la norme  $\mathcal{L}^2$  de la différence entre les analyses et l’analyse issue de l’algorithme exact.

Nous effectuons la minimisation. Le calcul des gradients de ces deux algorithmes est possible et est décrit dans l’annexe. Niveau qualité des résultats, nous retrouvons essentiellement les mêmes

analyses pour l’algorithme primal-dual. L’algorithme de régularisation entropique fournit un résultat plus lisse et d’amplitude moins bien capturée. Cependant cela revient à seulement 3% d’erreur sur la valeur de  $\mathcal{J}$ . Pour un  $\varepsilon$  plus petit, l’algorithme diverge numériquement. Nous pouvons penser que l’algorithme de régularisation entropique pour l’assimilation de données ne pourra pas du tout retrouver les détails.

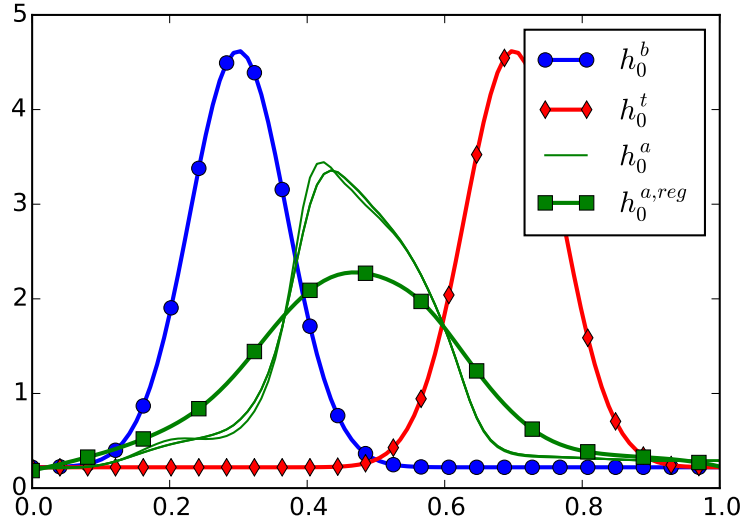


FIGURE 6.3 – Ébauche  $h_0^b$ , état vrai  $h_0^t$  qui a servi à faire les observations via un modèle de Shallow-Water, et les différentes analyses obtenues en fonction des différents algorithmes. Pour ne pas surcharger la légende  $h_0^a$  correspond aux analyses de tous les algorithmes utilisés (qui sont proches) sauf  $h_0^{a,reg}$  qui correspond à l’analyse issue de la régularisation d’entropie.

Au niveau temps de calcul, l’algorithme de régularisation entropique est très rapide, puisque un calcul de la distance est fait en moins d’une seconde. Comme attendu, l’algorithme primal-dual nécessite plus de calculs. Nous avons utilisé  $N = 100$  points de discrétisation, et il faut en moyenne une demi-seconde pour faire un calcul de  $\mathcal{W}_2$  pour le critère d’arrêt  $\varepsilon = 10^{-1}$  (cf. l’annexe A.5), et une dizaine de secondes pour le critère plus exigeant  $\varepsilon = 10^{-4}$ . Nous avons utilisé le fait que l’algorithme primal-dual est un algorithme itératif et qu’il peut être initialisé par le résultat d’un calcul précédent, ce qui divise par  $\sim 10$  le temps moyen de calcul quand on effectue plusieurs itérations de la minimisation de  $\mathcal{J}$ .

La complexité de l’algorithme primal-dual à nombre d’itérations fixé est en  $\mathcal{O}(N^{1+d} \ln(N))$  (résolution du laplacien) avec  $d$  la dimension spatiale. De plus, en suivant la tendance du temps de calcul de l’algorithme primal-dual implémenté pour [Farchi et al., 2016], où les temps en 1D et 2D sont comparés, s’il faut  $\sim 1s$  pour une grille de 100 points en 1D, alors il faudra  $\sim 100s$  pour un calcul sur une grille  $100 \times 100$ , ce qui dans certains cas peut poser problème.

Une possibilité pour améliorer les performances serait, lorsque l’on veut minimiser  $\mathcal{J}_{\mathcal{W}}$  avec des distances de Wasserstein calculées précisément à l’aide d’un algorithme primal-dual à  $\varepsilon$  faible, de commencer la minimisation en utilisant des distances de Wasserstein calculées grossièrement (avec des algorithmes de type régularisation entropique ou primal-dual à critères d’arrêt  $\varepsilon$  élevés), ce qui est rapide, avant d’affiner progressivement la précision du calcul numérique de  $\mathcal{W}_2$ .

Une autre optimisation possible peut être effectuée en calculant d’abord sur des grilles grossières avant d’affiner progressivement la grille, comme dans la méthode du 4D-Var multi-incrémental. Étant

donné que la distance de Wasserstein ne s'intéresse pas aux détails mais à la position des objets, comme nous l'avons vu, la calculer sur une grille grossière permettrait quand même de retrouver une masse de position intermédiaire.

Algorithme	Temps moyen	Erreur relative sur $\mathcal{J}_{exact}$	Erreur relative $\mathcal{L}_2$
Exact	0,02s	-	-
Primal-dual ( $\epsilon = 10^{-1}$ )	0,5s	$1,6 \cdot 10^{-2} \%$	4,2 %
Primal-dual ( $\epsilon = 10^{-4}$ )	10s	$2,8 \cdot 10^{-4} \%$	0,2 %
Régularisation ( $\epsilon = 10^{-2}$ )	0,3s	3,1%	29 %

TABLE 6.1 – Pour l'expérience d'assimilation de données et les différents algorithmes utilisés, on calcule le temps moyen du calcul de  $\mathcal{W}_2$ , ainsi que l'erreur relative avec l'algorithme de calcul exact, en terme de valeur de  $\mathcal{J}$  et de distance  $\mathcal{L}^2$  à la solution. Le maillage utilisé comporte 100 points.

### 6.2.3 Statistiques d'erreurs

Comme on l'a dit dans la section 1.1.2, on considère généralement l'erreur sur les données comme gaussienne et additive. Nous avons par exemple considéré que les observations diffèrent de l'état vrai par une telle erreur que nous avons écrite :

$$\mathbf{y}^{\text{obs}} = \mathcal{G}(\mathbf{x}_0^t) + \epsilon^{\text{obs}} \text{ avec } \epsilon^{\text{obs}} \sim \mathcal{N}(0, \mathbf{R})$$

avec  $\mathbf{R}$  une matrice qu'on considère diagonale par simplification. Choisir cette erreur implique naturellement un terme d'écart aux observations calculé à l'aide de la distance  $\mathcal{L}^2$ , en choisissant un estimateur *a posteriori*, et à l'aide du théorème de Bayes (cf. la démonstration des équivalences suivant la méthode 1.2.2). Dans la section 1.2.3 nous avons insisté sur le fait que l'on pouvait changer d'espace de représentation des données, et que c'est plutôt dans cet espace que l'erreur agit comme un bruit additif gaussien. Ainsi, si l'on note  $\mathcal{T}$  la transformation des données, on pourrait écrire

$$\mathcal{T}(\mathbf{y}^{\text{obs}}) = \mathcal{T}(\mathcal{G}(\mathbf{x}_0^t)) + \tilde{\epsilon}^{\text{obs}} \text{ avec } \tilde{\epsilon}^{\text{obs}} \sim \mathcal{N}(0, \tilde{\mathbf{R}})$$

ce qui amène à une autre fonction coût dont le terme d'écart à l'ébauche est

$$\|\mathcal{T}(\mathcal{G}(\mathbf{x}_0)) - \mathcal{T}(\mathbf{y}^{\text{obs}})\|_{\tilde{\mathbf{R}}}.$$

La question que l'on se pose est de savoir s'il y a un espace de représentation des données dans lequel si l'erreur est additive et gaussienne, alors la fonction coût associée sera écrite à l'aide de la distance de Wasserstein ?

Une piste de réflexion est d'utiliser les plans de transport pour décrire les erreurs de position. Nous considérons par exemple que les observations  $\mathbf{y}^{\text{obs}}$  diffèrent de l'état vrai  $\mathbf{x}_0^t$  à cause des erreurs de position. Nous écrivons alors

$$\mathcal{G}(\mathbf{x}_0^t) = (Id + \epsilon^{\text{obs}}) \# \mathbf{y}^{\text{obs}} \text{ avec } \epsilon^{\text{obs}} \sim \mathcal{N}(0, \mathbf{R}).$$

Cependant, ceci est mal défini, puisque pour  $\mathbf{y}^{\text{obs}}$  et  $\mathbf{x}_0^t$ , il existe une infinité de plans de transport  $T$  et donc de  $\epsilon^{\text{obs}}$  tel que  $T \# \mathbf{y}^{\text{obs}} = \mathcal{G}(\mathbf{x}_0^t)$ . Notamment, on pourrait avoir  $\epsilon^{\text{obs}}$  aussi grand que l'on veut mais tel que  $(Id + \epsilon^{\text{obs}}) \# \mathcal{G}(\mathbf{x}_0^t)$  soit proche de  $\mathcal{G}(\mathbf{x}_0^t)$  : par exemple en prenant un  $\epsilon^{\text{obs}}$  à divergence nulle, qui fait tourner localement un point sur lui-même sans le déplacer.

Définir de manière bijective une transformation entre l'espace des états et l'espace des plans de transport, ou définir de manière unique l'erreur sur la position, est possible en considérant seulement certains types de plans de transport : les plans de transport optimaux. L'application bijective  $\mathcal{T}$  est alors :

$$\mathcal{T}(\mathbf{y}) = \text{Plan de transport optimal entre } \mathbf{y}^{\text{obs}} \text{ et } \mathbf{y}.$$

Les plans de transport sont de la forme  $T(x) = x + \nabla\Phi(x)$ . Nous écrivons que  $\mathbf{y}^{\text{obs}}$  est généré à partir de  $\mathbf{x}_0^t$  avec de l'erreur de position, par la formule :

$$\mathcal{G}(\mathbf{x}_0^t) = (Id + \nabla\Phi^{\text{obs}})\#\mathbf{y}^{\text{obs}} \quad \text{avec } \nabla\Phi^{\text{obs}} \sim \mathcal{N}(0, \mathbf{R}). \quad (6.2.2)$$

Ici,  $\mathbf{R}$  est une fonction de covariance d'erreur, que nous faisons dépendre de  $x$  et de  $y$  :  $\mathbf{R}(x, y)$  définie comme

$$\forall x, y, \quad \mathbf{R}(x, y) := \mathbb{E}(\nabla\Phi^{\text{obs}}(x)^t \nabla\Phi^{\text{obs}}(y)).$$

La matrice de covariance d'erreur classique est obtenue en discrétisant. Ainsi, le terme dans la fonction coût associée à cette erreur est :

$$\|\nabla\Phi^{\text{obs}}\|_{\mathbf{R}}^2 = \iint_{\Omega \times \Omega} \nabla\Phi^{\text{obs}}(x)^t \mathbf{R}^{-1}(x, y) \nabla\Phi^{\text{obs}}(y) \, dx dy$$

L'approximation classique d'assimilation de données est de prendre  $\mathbf{R}$  diagonale. Dans le cas particulier où

$$\mathbf{R}(x, y) = \frac{\sigma_{\mathbf{R}}^2}{\mathbf{y}^{\text{obs}}} \delta_{x=y},$$

alors on peut écrire pour  $\Phi^{\text{obs}}$  le potentiel de transport entre  $\mathbf{y}^{\text{obs}}$  et un  $\mathcal{G}(\mathbf{x}_0)$ ,

$$\begin{aligned} \|\nabla\Phi^{\text{obs}}\|_{\mathbf{R}}^2 &= \frac{1}{\sigma_{\mathbf{R}}^2} \int_{\Omega} \mathbf{y}^{\text{obs}} |\nabla\Phi^{\text{obs}}(x)|^2 \, dx \\ &= \frac{1}{\sigma_{\mathbf{R}}^2} \mathcal{W}_2(\mathcal{G}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 \end{aligned}$$

car justement  $x + \nabla\Phi^{\text{obs}}$  est le plan de transport optimal entre  $\mathbf{y}^{\text{obs}}$  et  $\mathcal{G}(\mathbf{x}_0)$ . Nous obtenons donc un terme de la fonction coût écrite avec la distance de Wasserstein.

Nous pouvons faire de même pour l'ébauche et écrire :

$$\mathbf{x}_0^t = (Id + \nabla\Phi^{\text{obs}})\#\mathbf{x}_0^b \quad \text{avec } \nabla\Phi^b \sim \mathcal{N}(0, \mathbf{B})$$

qui vaut  $\frac{1}{\sigma_{\mathbf{B}}^2} \mathcal{W}_2(\mathbf{x}_0, \mathbf{x}_0^b)^2$  dans le cas où  $\mathbf{B} = \frac{\sigma_{\mathbf{B}}^2}{\mathbf{x}_0^b(x)} \delta_{x=y}$ .

Dans ces deux cas de matrices  $\mathbf{R}$  et  $\mathbf{B}$ , nous retrouvons  $\mathcal{J}_{\mathcal{W}}$  puisqu'alors nous avons :

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{\sigma_{\mathbf{R}}^2} \mathcal{W}_2(\mathcal{H}(\mathbf{x}_0), \mathbf{y}^{\text{obs}})^2 + \frac{1}{\sigma_{\mathbf{B}}^2} \mathcal{W}_2(\mathbf{x}_0, \mathbf{x}_0^b)^2.$$

Par contre dans un cas plus général, par exemple pour une fonction diagonale non constante

$$\mathbf{B}(x, y) = \frac{1}{\mathbf{s}(x)\mathbf{x}_0^b(x)}\delta_{x=y},$$

avec  $\mathbf{s}(x)$  une fonction positive, nous ne semblons pas pouvoir retrouver  $\mathcal{J}_{\mathcal{W}}$ . Il existe des distances de Wasserstein pondérées, comme le transport optimal anisotrope [Hug et al., 2015], définis pour une fonction  $\mathbf{s}(x)$  par

$$\mathcal{W}_{\mathbf{s}}(\rho_0, \rho_1) = \inf_{(\rho, v) \in \mathcal{C}(\rho_0, \rho_1)} \iint_{[0,1] \times \Omega} \rho(t, x) v(t, x)^t \mathbf{s}(x) v(t, x) dt dx$$

mais dont les solutions ne semblent pas s'écrire à l'aide d'un plan de transport (il n'y a pas nécessairement de problème de Monge lié).

Le problème est que le terme lié dans la fonction coût est, pour  $\Phi$  un potentiel de transport entre l'ébauche et un  $\mathbf{x}_0$ ,

$$\|\nabla\Phi\|_{\mathbf{B}} = \int_{\Omega} \mathbf{x}_0^b(x) \nabla\Phi(x)^t \mathbf{s}(x) \nabla\Phi(x) dx$$

mais ceci n'est en fait pas égal à  $\mathcal{W}_{\mathbf{s}}(\mathbf{x}_0^b, \mathbf{x}_0)$ . Nous n'avons pas pu trouver comment contourner le problème, c'est-à-dire à savoir comment faire lorsque  $\mathbf{B}$  est à diagonale non constante.

Ce transport optimal anisotrope permet notamment de calculer des distances de Wasserstein dans le cas où l'espace contient des obstacles que le transport doit contourner : en mettant  $\mathbf{s}(x) = +\infty$  au niveau des obstacles, on trouve un transport optimal l'évitant. Une application est l'utilisation du transport optimal en assimilation de données pour l'océanographie, en intégrant les côtes par exemple, qui sont des obstacles, cf. la figure 6.4.

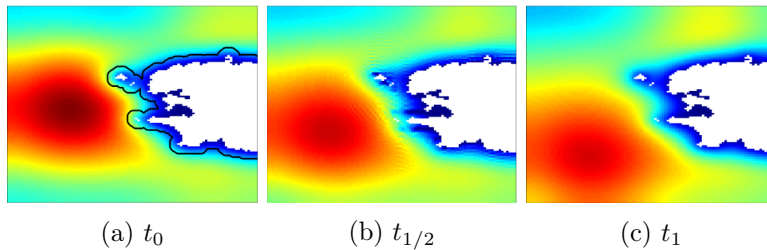


FIGURE 6.4 – Transport optimal entre la figure de gauche (a) à  $t = 0$  et la figure de droite (b) à  $t = 1$ , en prenant en compte la côte bretonne, ce qui donne l'interpolation (c) à  $t = 1/2$ . Source : [Papadakis et al., 2014].

De même que pour  $\mathcal{W}_{\mathbf{s}}$ , pour une matrice générale (non-diagonale)  $\mathbf{B}(x, y) = \frac{1}{\sqrt{\rho_0(x)\rho_1(y)}}\mathbf{A}^{-1}(x, y)$ , nous n'avons pas établi la valeur de la fonction coût, même si nous pouvons penser à vouloir relier à un terme s'écrivant :

$$\mathcal{W}_{\mathbf{B}}(\rho_0, \rho_1)^2 := \iiint_{[0,1] \times \Omega \times \Omega} \sqrt{\rho(t, x)} v(t, x)^t \mathbf{A}(x, y) v(t, y) \sqrt{\rho(t, y)} dt dx dy. \quad (6.2.3)$$

Nous n'avons donc pas totalement réglé le problème de savoir comment écrire les erreurs pour que la fonction coût liée soit de type Wasserstein. En écrivant les erreurs de position à l'aide d'un plan de transport, seul le cas où les matrices  $\mathbf{B}$  et  $\mathbf{R}$  sont à diagonale inversement proportionnelle à  $\mathbf{x}_0^b$  et à  $\mathbf{y}^{\text{obs}}$  donne une fonction coût de Wasserstein. Ce problème reste donc ouvert et pourrait faire le lien avec des distances de Wasserstein anisotropes ou comme dans (6.2.3).

### 6.2.4 Occultations

Nous avons fait l'hypothèse dans nos expériences que les observations étaient totales, c'est-à-dire la mesure d'une variable sur tout  $\Omega$ , c'est-à-dire que  $\Omega = \Omega^{\text{obs}}$ . Ceci sur-détermine le problème puisqu'il y a plus d'observations que d'inconnues. Les observations sont rarement faites sur tout  $\Omega$  et nous étudions comment elles sont utilisées en assimilation de données, pour s'intéresser à une version écrite avec la distance de Wasserstein. Ces observations peuvent être :

- Des captures de la surface, comme en océanographie où la température de la surface de l'océan est mesurée par les satellites. Ainsi, dans une assimilation d'un modèle océanique, l'opérateur  $\mathcal{H}$  restreint tout l'océan à sa surface.
- Des intégrations selon une direction, comme en météorologie où les satellites intègrent l'humidité selon la verticale. Cela crée une observation 2D d'un système 3D.
- Des observations sur une zone seulement : comme en météorologie où un satellite géostationnaire observe une partie de la Terre seulement.
- Des occultations peuvent apparaître, tels des nuages sur une image satellite.

Dans tous les cas l'opérateur  $\mathcal{H}$  restreint à une zone de l'espace. Nous étudions les conséquences que ceci peut avoir dans une assimilation de 3D-Var. La fonction coût euclidienne s'écrit :

$$\mathcal{J}(h) = \|\mathbf{H}h - y^{\text{obs}}\|_2^2 + \|h - h^b\|_2^2$$

avec  $\mathbf{H}$  l'opérateur linéaire de restriction de  $\Omega$  à  $\Omega^{\text{obs}}$  (la zone d'observation incluse dans  $\Omega$ , elle peut être une zone de  $\Omega$  ou un de ses bords par exemple). Le minimiseur  $h$  vérifie les conditions d'Euler-Lagrange :

$$\mathbf{H}^*(\mathbf{H}h - y^{\text{obs}}) + h - h^b = 0$$

où  $\mathbf{H}^*$  est l'opérateur adjoint à  $\mathbf{H}$ , et égal à l'identité sur  $\Omega^{\text{obs}}$  et est nul ailleurs. Ainsi, le minimiseur  $h$  est égal à

$$h = \begin{cases} \frac{y^{\text{obs}} + h^b}{2} & \text{dans } \Omega^{\text{obs}} \\ h^b & \text{dans } \Omega \setminus \Omega^{\text{obs}}. \end{cases}$$

Une conséquence notamment est une discontinuité sur la frontière entre  $\Omega^{\text{obs}}$  et  $\Omega \setminus \Omega^{\text{obs}}$ .

Une des solutions pour éviter ce problème est d'incorporer une matrice de covariance d'erreur  $\mathbf{B}$  pour pondérer le terme d'écart à l'ébauche et qui va avoir pour effet de lisser les données [Blayo et al., 2015, sec. 8.2.1]. En effet, l'opérateur  $\mathbf{B}$  peut être vu comme un opérateur de diffusion. Grâce à cette matrice, la solution devient continue.

Dans le cas de la distance de Wasserstein, nous écrivons la fonction coût comme

$$\mathcal{J}(h) = \mathcal{W}_2(\mathbf{H}h, y^{\text{obs}})^2 + \mathcal{W}_2(h, h^b)^2$$

Cette écriture suppose que  $\mathbf{H}h$  et  $y^{\text{obs}}$  sont de même masse (quel que soit  $h \in \mathcal{P}(\Omega)$ ). Si ce n'est pas le cas, on utilisera plutôt la distance  $WF$ . Nous affichons sur la figure 6.5 deux exemples attendus de la minimisation. Ces exemples ont une partie de l'espace non observé. Ces exemples nous font penser que la plupart du temps, le minimiseur se situera sous forme de Dirac au bord du domaine observé  $\Omega^{\text{obs}}$ .

Si nous faisons un parallèle entre problèmes de discontinuités (présentes sur le bord de  $\Omega^{\text{obs}}$ ) dans le cas  $\mathcal{L}^2$  et Diracs dans le cas  $\mathcal{W}_2$ , alors on peut supposer qu'ajouter une matrice de covariance d'erreur  $\mathbf{B}$  du même type que pour  $\mathcal{L}^2$  permettrait de résoudre le problème. La distance de Wasserstein avec la matrice  $\mathbf{B}$  s'écrirait comme dans l'équation (6.2.3).

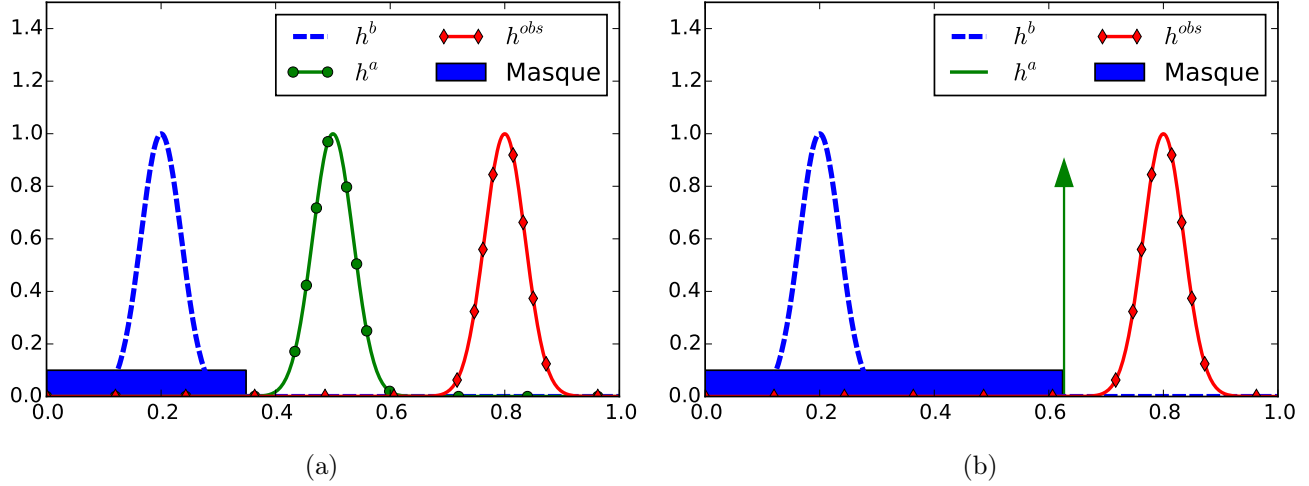


FIGURE 6.5 – Minimisation de  $\mathcal{J}(h) = \mathcal{W}_2(\mathbf{H}h, h^{\text{obs}})^2 + \mathcal{W}_2(h, h^b)^2$  dans deux cas : dans le premier (a) où seule la partie  $[0, 0.35]$  est occultée par le masque tracé. Seul la partie  $[0.35, 1]$  est observée. Le minimiseur  $h$  coïncide avec celui dans le cas où il n’y aurait pas de masque, étant donné que  $\mathbf{H}h = h$ . Dans (b), où le masque est plus grand ( $[0, 0.6]$ ), on remarque que  $\mathbf{H}h = 0$  ce qui est très pénalisant dans le terme d’écart aux observations. Plutôt, nous faisons l’hypothèse que le minimiseur aura tendance à se coller le long de la frontière de  $\Omega^{\text{obs}}$ .

### 6.2.5 Assimilation de données par filtrage et par la distance de Wasserstein

Nous avons évoqué dans le chapitre 1 l’assimilation de données par filtrage, basée sur le filtre de Kalman. Nous montrons ici que cette assimilation de donnée est possible en utilisant une distance de Wasserstein et donc en jouant sur les positions des états. Si nous reprenons la méthode 1.2.5 du filtre de Kalman, celui-ci se décompose en trois étapes :

**Prévision** Au temps  $t_i$  nous supposons avoir estimé l’état  $\mathbf{x}_i$ . Cet état est intégré jusqu’au temps  $t_{i+1}$  *via* le modèle, mais sans prendre en compte l’observation  $\mathbf{y}_{i+1}^{\text{obs}}$ . Nous définissons la prévision :

$$\hat{\mathbf{x}}_{i+1|i} = \mathcal{M}_i(\mathbf{x}_i).$$

**Correction** Cet état est corrigé en prenant en compte l’observation  $\mathbf{y}_{i+1}^{\text{obs}}$ , avec une formule du BLUE. Mais au lieu d’utiliser des distances euclidiennes comme dans l’équation (1.2.6), nous pouvons définir la correction comme :

$$\hat{\mathbf{x}}_{i+1} = \arg \min [\mathcal{W}_2(\mathcal{H}_{i+1}(\mathbf{x}_{i+1}), \mathbf{y}_{i+1}^{\text{obs}})^2 + \mathcal{W}_2(\mathbf{x}_{i+1}, \hat{\mathbf{x}}_{i+1})^2]$$

ou, si l’on veut ajouter des matrices de covariance d’erreur,

$$\hat{\mathbf{x}}_{i+1} = \arg \min [\mathcal{W}_{\mathbf{R}_{i+1}}(\mathcal{H}_{i+1}(\mathbf{x}_{i+1}), \mathbf{y}_{i+1}^{\text{obs}})^2 + \mathcal{W}_{\mathbf{B}_{i+1}}(\mathbf{x}_{i+1}, \hat{\mathbf{x}}_{i+1})^2]$$

avec  $\mathbf{R}_{i+1}$  et  $\mathbf{B}_{i+1}$  les matrices de covariance d’erreur associées respectivement à l’observation  $\mathbf{y}_{i+1}^{\text{obs}}$  et l’ébauche. Nous avons défini  $\mathcal{W}_{\mathbf{R}_{i+1}}$  et  $\mathcal{W}_{\mathbf{B}_{i+1}}$  comme dans (6.2.3).

**Calcul des matrices de covariance d’erreur.** Le filtre de Kalman passe ensuite par une évaluation de la matrice  $\mathbf{B}_{i+1}$  notamment. Dans le cas de distances euclidiennes et de modèles linéaires, cette matrice est calculable. Pour le cas de la distance de Wasserstein et de modèles linéaires, cela reste encore un problème ouvert et des travaux supplémentaires de recherche pourront s’axer sur ces matrices.





# Annexe A

## Résolution numérique du transport optimal

### Contenu du chapitre

A.1 Cas uni-dimensionnel . . . . .	141
A.2 Versions primale, duale et primale-duale du transport optimal . . . . .	142
A.3 Résolution du problème primal par descente de gradient . . . . .	144
A.4 Algorithme pour la version duale . . . . .	146
A.5 Résolution du problème primal-dual . . . . .	147
A.6 Régularisation d'entropie . . . . .	151
A.7 Algorithme pour un calcul numérique de la distance WF. . . . .	154

Nous nous intéressons dans cette annexe aux calculs numériques du transport optimal. En 1D il existe une formule exacte que nous détaillerons. Il existe également d'autres algorithmes qui peuvent être étendus en dimension supérieure, mais qui peuvent aussi s'adapter pour le calcul de la distance  $WF$  en 1D. Nous étudierons en particulier l'algorithme primal-dual, assez lent mais précis, et le transport optimal par régularisation d'entropie, rapide mais qui est seulement une approximation.

### A.1 Cas uni-dimensionnel

Évoquons le calcul uni-dimensionnel du transport optimal. On considère  $\Omega = [0, 1]$ . Il existe une formule pour retrouver la distance de Wasserstein.

#### **Théorème A.1.1** (*Formule du plan de transport optimal unidimensionnel*)

Soit deux densités  $\rho_0, \rho_1$  de  $\mathcal{P}(\Omega)$ , le plan de transport optimal  $T(x)$  est croissant et vérifie

$$\forall x \in [0, 1], \quad F_1(T(x)) = F_0(x) \tag{A.1.1}$$

avec  $F_i(x)$  la fonction de distribution de  $\rho_i$  ( $i = 0$  ou  $1$ ),

$$\forall x, F_i(x) = \int_0^x \rho_i(y) dy.$$

**Démonstration :** Le plan de transport optimal  $T$  est de la forme  $T = \Psi'$  avec  $\Psi$  convexe, comme le veut le Théorème de Brenier 2.3.1. Ceci implique que le plan de transport  $T$  est croissant. De plus,  $T$  vérifie  $T\#\rho_0 = \rho_1$ , ce qui s'écrit

$$\forall y \in \Omega, \quad \rho_1(T(y))T'(y) = \rho_0(y)$$

et qui redonne la formule (A.1.1) en intégrant entre 0 et  $x$ . ■

Pour calculer  $T(x)$  il suffit juste de résoudre l'équation (A.1.1). Numériquement, on effectue simplement une dichotomie pour trouver le zéro de la fonction  $F_1(T) - F_0(x)$ , pour chaque  $x$ . La distance de Wasserstein est ensuite donnée par

$$\mathcal{W}_2^2(\rho_0, \rho_1) = \int_0^1 \rho_0(x) |T(x) - x|^2 dx.$$

En deux dimensions il n'existe pas de formule pour calculer directement la distance de Wasserstein. Plusieurs algorithmes ont cependant été développés.

## A.2 Versions primale, duale et primale-duale du transport optimal

On a défini la distance de Wasserstein notamment par sa formulation Benamou-Brenier qui est :

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \inf_{(\rho, v) \in C(\rho_0, \rho_1)} \frac{1}{2} \iint \rho |v|^2 dx dt$$

$$\text{où } \tilde{C}(\rho_0, \rho_1) := \left\{ (\rho, v) \text{ tel que } \begin{array}{l} \partial_t \rho + \operatorname{div}(\rho v) = 0, \\ \rho(t=0) = \rho_0, \rho(t=1) = \rho_1, \\ \rho v \cdot \vec{n} = 0 \text{ sur } [0, 1] \times \partial\Omega \end{array} \right\}. \quad (\text{A.2.1})$$

On peut remarquer que ces contraintes sont non-linéaires et la fonction n'est pas convexe, c'est donc délicat à résoudre. Mais en effectuant le changement de variable  $m := \rho v$ , le problème devient

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \inf_{(\rho, m) \in C'(\rho_0, \rho_1)} \iint \frac{|m|^2}{2\rho} dt dx$$

et la convention  $\frac{|m|^2}{2\rho} = 0$  si  $\rho = 0$ . L'ensemble de contraintes devient lui

$$C(\rho_0, \rho_1) := \left\{ (\rho, m) \text{ tel que } \begin{array}{l} \partial_t \rho + \operatorname{div}(m) = 0, \\ \rho(t=0) = \rho_0, \rho(t=1) = \rho_1, \\ m \cdot \vec{n} = 0 \text{ sur } [0, 1] \times \partial\Omega \end{array} \right\}. \quad (\text{A.2.2})$$

Cet ensemble des contraintes est convexe et la fonction coût est convexe, on peut donc résoudre le problème à l'aide des méthodes habituelles. On pose  $U = (\rho, m)$ ,

$$J(U) := \iint_{[0,1] \times \Omega} \frac{m^2}{2\rho} dt dx$$

et le problème de transport optimal consiste à résoudre

$$\min_U J(U) + \iota_C(U).$$

La fonction  $\iota_C$  est la fonction caractéristique de  $C$ , valant 0 sur  $C$  et  $+\infty$  ailleurs. On peut résoudre ce problème par gradient projeté par exemple, voir la section A.3.

On verra qu'il sera plus efficace de reformuler ce problème dans sa version primale-duale via la théorie classique des fonctions convexes et de la dualité de Fenchel-Rockafellar. En effet, comme  $J$  est convexe, elle est égale à deux fois sa transformée de Fenchel-Legendre,  $J = J^{**}$ , où l'étoile  $*$  définit la transformée de Fenchel-Legendre : par exemple,  $J^*(V)$  est la transformée de  $J(U)$  définie par

$$J^*(V) = \sup_U (\langle U, V \rangle - J(U)).$$

La variable duale  $V = (a, b)$  est la variable de  $J^*$ . Les calculs donnent que  $J^*(V)$  s'écrit juste comme la fonction caractéristique d'un ensemble,

$$J^*(V) = \iota_K(V) \quad \text{avec} \quad K = \left\{ (a, b) : a + \frac{b^2}{2} \leq 0 \right\}$$

On introduit donc cette transformée dans le problème de transport optimal, qui s'écrit via l'égalité  $J = J^{**}$  :

$$\begin{aligned} \min_U J(U) + \iota_C(U) &= \min_U J^{**}(U) + \iota_C(U) = \min_U \max_V \langle U, V \rangle - J^*(V) + \iota_C(U) \\ &= \min_U \max_V \langle U, V \rangle - \iota_K(V) + \iota_C(U) \end{aligned}$$

Ce dernier min max est la version **primale-duale** du transport optimal.

On aurait pu passer à la version complètement duale du transport optimal, en écrivant le  $\max_U \langle U, V \rangle + \iota_C(U)$  comme la transformée de Fenchel-Legendre de  $\iota_C$ ,

$$\begin{aligned} \min_U \max_V \langle U, V \rangle - \iota_K(V) + \iota_C(U) &= \max_V -\iota_K(V) + \min_U \langle U, V \rangle + \iota_C(U) \\ &= \max_V -\iota_K(V) - \iota_C^*(-V) \end{aligned}$$

Ceci est la version **duale** du transport optimal. On calcule que

$$\begin{aligned} \iota_C^*(-V) &= \int_{\Omega} [\rho_1 \Phi(1, x) - \rho_0 \Phi(0, x)] dx \quad \text{si } V = \nabla_{t,x} \Phi \\ &= +\infty \quad \text{sinon.} \end{aligned}$$

On peut ainsi rapprocher la version duale de la version duale de Kantorovich.

Numériquement, il existe des algorithmes différents qui se basent chacun sur l'une des trois versions (primale, primale-duale et duale). Les différents algorithmes créent des suites de variables  $U^n(t, x)$  et/ou  $V^n(t, x)$  qui convergent vers l'optimum.

La grille de discrétisation est spatio-temporelle : les  $U^n$  et  $V^n$  sont évalués sur des points  $(t_n, x_i)_{n,i}$ . Nous aurions pu prendre des grilles centrées classiques, mais il semble plus efficace de choisir des grilles décalées, comme suggéré dans [Papadakis et al., 2014]. La variable  $\rho$  de  $U$  vit sur une grille décalée en espace  $\mathcal{G}_d^x$ , la variable  $m$  sur une grille décalée en temps  $\mathcal{G}_d^t$  et les variables  $a$  et  $b$  de  $V$  sont toutes les deux décalées en espace-temps  $\mathcal{G}_c$ , comme sur la figure A.1.

On note  $\mathcal{U}$  les fonctions définies sur  $\mathcal{G}_d^x \times \mathcal{G}_d^t$  et  $\mathcal{V}$  les fonctions définies sur  $\mathcal{G}_c \times \mathcal{G}_c$ . Enfin, on crée l'opérateur  $I : \mathcal{U} \rightarrow \mathcal{V}$  qui interpole  $U$  sur la grille centrée,

$$(IU)_{i+\frac{1}{2}, j+\frac{1}{2}} = \frac{1}{2} \begin{pmatrix} \rho_{i+1, j+\frac{1}{2}} + \rho_{i, j+\frac{1}{2}} \\ m_{i+\frac{1}{2}, j+1} + m_{i+\frac{1}{2}, j} \end{pmatrix}$$

L'adjoint est  $I^* : \mathcal{V} \rightarrow \mathcal{U}$ . De plus, la norme sur  $\mathcal{V}$  et la norme sur  $\mathcal{U}$  sont respectivement

$$\begin{aligned} \|V\|^2 &= \sum_{i,j} a_{i+\frac{1}{2}, j+\frac{1}{2}}^2 + b_{i+\frac{1}{2}, j+\frac{1}{2}}^2 \\ \|U\|^2 &= \iint (\rho^2 + m^2) dt dx = \sum_{i,j} \rho_{i, j+\frac{1}{2}}^2 + m_{i+\frac{1}{2}, j}^2 - \frac{1}{2} Tr(U^2) \end{aligned}$$

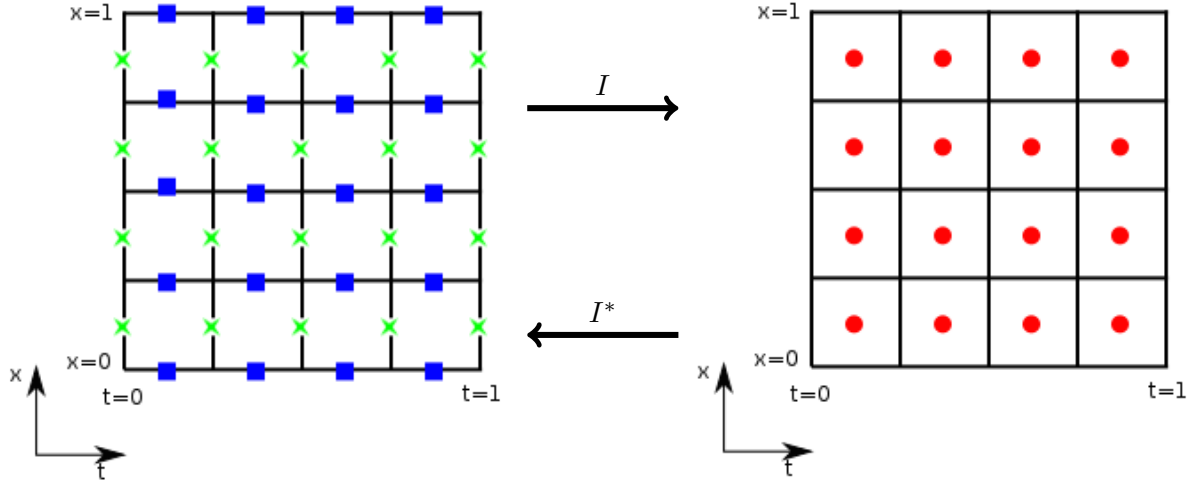


FIGURE A.1 – À gauche : les croix vertes représentent les points de la grille décalée en temps  $\mathcal{G}_d^t$ , et les carrés bleus les points de la grille décalée en espace  $\mathcal{G}_d^x$ . À gauche, les points rouges sont sur la grille centrée  $\mathcal{G}_c$ .

avec  $Tr(U^2)$  l'intégrale de  $(\rho^2, m^2)$  sur le bord du domaine (la trace). On ajoute ce terme de trace dans la définition de la norme pour  $\mathcal{U}$  pour avoir une discrétisation trapézoïdale de l'intégrale spatio-temporelle. Si l'on avait utilisé des rectangles centrés pour la discrétisation de l'intégrale, cela définirait seulement une pseudo-norme (on pourrait avoir  $\|U\| = 0$  mais  $U \neq 0$ ). Enfin, avec ces définitions, on obtient la propriété intéressante d'adjoint de l'opérateur divergence

$$\text{div}_{t,x}^* = -\nabla_{t,x} : \mathcal{G}_c \rightarrow \mathcal{U},$$

dans le sens où  $\langle \text{div}_{t,x}(U), \Phi \rangle_{\mathcal{G}_c} = -\langle \nabla_{t,x} \Phi, U \rangle_{\mathcal{U}}$ , et avec  $\nabla_{t,x}$  l'opérateur gradient défini par différences finies centrées.

Pour en revenir au problème de transport optimal, la fonction coût qu'on minimise est

$$J(IU) = \sum_{i,j} \frac{\tilde{m}_{i+\frac{1}{2},j+\frac{1}{2}}^2}{2\tilde{\rho}_{i+\frac{1}{2},j+\frac{1}{2}}}$$

avec  $(\tilde{\rho}, \tilde{m}) = I(\rho, m)$ . Pour résumer, et en incluant  $I$  dans les calculs précédemment faits, les trois versions primale, duale et primale-duale du problème s'écrivent respectivement

$$\mathcal{W}_2(\rho_0, \rho_1)^2 = \min_U J(IU) + \iota_C(U) \quad (\text{P})$$

$$= \min_U \max_V \langle IU, V \rangle - \iota_K(V) + \iota_C(U) \quad (\text{PD})$$

$$= \max_V -\iota_K(V) - \iota_C^*(-I^*V) \quad (\text{D})$$

Ces trois formulations donnent chacune lieu à des algorithmes de résolution différents. On va les présenter ci-dessous.

### A.3 Résolution du problème primal par descente de gradient

On résout le problème primal (P) par méthode de gradient projeté :

**Algorithme A.3.1 (Gradient projeté)**

1. calculer le gradient de  $J$ ,  $\text{grad}J$ , valant

$$\text{grad} J(IU) = I^* \begin{pmatrix} -\frac{\tilde{m}^2}{2\tilde{\rho}} \\ \frac{\tilde{m}}{\tilde{\rho}} \end{pmatrix}$$

avec  $I^*$  l'opérateur adjoint de  $I$  et  $(\tilde{\rho}, \tilde{m}) = IU$  ;

2. descendre dans la direction du gradient, en projetant dans l'espace des contraintes  $C$ , via l'itération

$$U^{n+1} = P_C(U^n - \alpha \text{grad} J(IU^n)).$$

Étant donné que l'espace  $C$  est affine on peut projeter le gradient dans l'espace tangent à  $C$ ,  $TC$  qui est défini comme  $C$  mais avec les constantes égales à 0

$$TC = \left\{ (\eta, \mu) \text{ tel que } \begin{array}{l} \partial_t \eta + \text{div}(\mu) = 0, \\ \eta(t=0) = 0, \eta(t=1) = 0, \\ \mu \cdot \vec{n} = 0 \text{ sur } [0, 1] \times \partial\Omega \end{array} \right\}.$$

L'itération donne

$$U^{n+1} = U^n - \alpha P_{TC}(\text{grad} J(U^n)).$$

La valeur de la projection sur  $TC$  est explicitée par la proposition A.3.2 ci-dessous.

**Proposition A.3.2**

La projection de  $g = (\eta, \mu) \notin TC$  sur  $TC$  est définie par

$$P_{TC}(g) = g - \nabla\Phi$$

avec  $\Phi$  résolvant le Laplacien aux conditions aux bords de Neumann,

$$\begin{cases} \Delta\Phi = \text{div}(g) \\ \nabla\Phi \cdot \vec{n} = (\eta(0), \eta(1), \mu(x=0), \mu(x=1)) \end{cases}$$

**Démonstration :** Soit  $g = (\eta, \mu) \notin TC$ , pour tout  $g' \in TC$  on a

$$\langle g - P_{TC}(g), g' - P_{TC}(g) \rangle = 0.$$

Ce qui se traduit (l'espace  $TC$  étant linéaire), en remplaçant  $g'$  par  $g' - P_{TC}(g) = g''$ , par

$$\forall g'' \in TC, \quad \langle g - P_{TC}(g), g'' \rangle = 0.$$

d'où  $g - P_{TC}(g) \in \ker(\text{div})^\perp = \text{Im}(\text{div}^* = -\nabla)$ , donc  $g - P_{TC}(g) = \nabla\Phi$ . Pour trouver la valeur de  $\Phi$ , il suffit d'appliquer l'opérateur divergence et d'inverser le Laplacien. ■

Ceci est la méthode du gradient projeté. Nous l'avons comparé avec la méthode du gradient conjugué projeté, qui donne de meilleurs résultats. En effet, dans la figure A.2 nous comparons les résultats de convergence des deux algorithmes de descente de gradient et de descente de gradient conjugué, en une dimension. Pour le gradient conjugué, nous utilisons la variante de Hestenes-Stiefel, avec réinitialisation de la direction de descente tous les 100 itérations, une des meilleures combinaisons trouvée [Nocedal et Wright, 2006].

Pour une méthode de gradient conjugué, les résultats sont bien meilleurs. Nous remarquons aussi que la descente de gradient simple n'arrive pas à converger. Ceci est dû à la convexité de  $J$  ; même si dans [Henry, 2016] on peut trouver un résultat qui dit que  $J$  est strictement convexe dans  $C$ ,  $J$  a peu de chance d'être fortement convexe dans  $C$ , même autour de l'optimum.

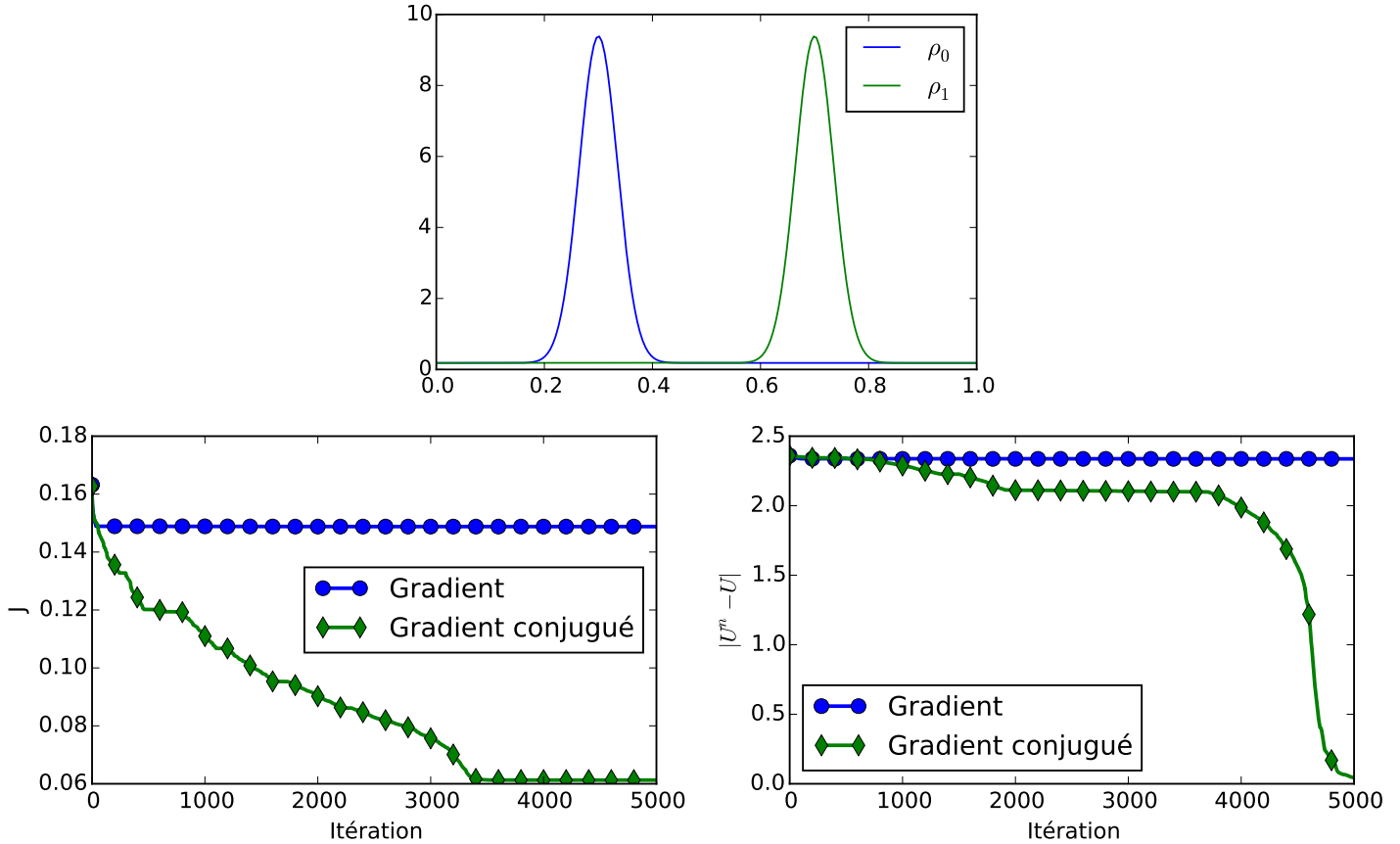


FIGURE A.2 – Comparaison des algorithmes de descente de gradient simple et de gradient conjugué, pour le calcul du transport optimal entre deux densités  $\rho_0$  et  $\rho_1$  (en haut). Le gradient conjugué utilise la version "Hestenes-Stiefel" avec réinitialisation de la direction de descente toutes les 100 itérations. En bas à gauche est affichée la valeur de  $J$  en fonction des itérations. En bas à droite est tracée la valeur de  $\|U^n - U\|$  en fonction de l'itération  $n$ , où  $U$  est la valeur optimale, obtenue après convergence d'un autre algorithme plus performant (voir A.5).

Ces calculs sont classiques mais peu encourageants : le calcul de la distance est assez long. Il faut environ 3000 itérations (donc 3000 projections donc 3000 inversions de Laplacien sur la grille spatio-temporelle) avant d'avoir la convergence. D'autres algorithmes prélevés du traitement d'image ont été créés, se basant d'abord sur la version duale puis sur la version primale-duale du transport optimal.

#### A.4 Algorithme pour la version duale

Le premier réel algorithme pour résoudre le problème de transport optimal entre deux densités a été développé en 2000 et porte le nom d'algorithme de Benamou-Brenier, qui repose sur

un Lagrangien augmenté et permet de résoudre l'opposé de la version duale (D) du problème de transport

$$-(D): \min_V \iota_K(V) + \iota_C^*(-I^*V)$$

avec

$$\begin{aligned} \iota_C^*(-I^*V) &= \int_{\Omega} [\rho_1 \Phi(1, x) - \rho_0 \Phi(0, x)] dx & \text{si } -I^*V = \nabla_{t,x} \Phi \\ &= +\infty & \text{sinon} \end{aligned}$$

Pour résoudre ce problème, étant donnée la formulation de  $\iota_C^*$ , on ajoute la contrainte  $I^*V = \nabla \Phi$  et donc on ajoute un multiplicateur de Lagrange associé à cette contrainte, et le problème à résoudre est le point-selle suivant, augmenté par un terme quadratique pour des raisons calculatoires,

$$\inf_{V, \Phi} \sup_U \underbrace{\iota_C^*(-\nabla \Phi) + \iota_K(V) + \langle U, \nabla_{t,x} \Phi - I^*V \rangle + \frac{r}{2} \|\nabla \Phi - I^*V\|^2}_{\mathcal{L}_r(U, V, \Phi)}$$

Ce type problème est résolu de manière assez performante par les algorithmes développés par Fortin et Glowinski dans [Fortin et Glowinski, 1983], et est détaillé pour le problème du transport optimal dans [Benamou et Brenier, 2000]. L'algorithme crée une suite  $(U^n, V^n, \Phi^n)$  qui converge vers le point-selle du Lagrangien, une démonstration étant présentée dans [Hug, 2016]. L'algorithme se divise en trois étapes, une première est une projection sur  $C$ , une deuxième étant une projection sur  $K$ , et une troisième de simple incrément sur  $U$ . On ne détaille pas plus l'algorithme puisque récemment un autre, lui ressemblant, se montre plus performant. C'est un algorithme se basant sur la version primale-duale (PD).

## A.5 Résolution du problème primal-dual

Depuis, on peut résoudre de manière beaucoup plus efficace le problème du transport optimal, par exemple par l'algorithme primal-dual développé par Chambolle et Pock dans [Chambolle et Pock, 2011] et récemment mis en œuvre pour le transport optimal [Papadakis et al., 2014]. Ces algorithmes primaux-duaux permettent de résoudre des points-selles de fonctions qui s'écrivent

$$\mathcal{L}(x, y) = \langle Kx, y \rangle + G(x) - F^*(y). \quad (\text{A.5.1})$$

avec  $K$  linéaire,  $G$  et  $F^*$  des fonctions convexes et semi-continues inférieurement. Ce type d'algorithme s'utilise beaucoup en imagerie (par exemple en restauration d'image [Esser, 2010]) puisqu'en imagerie on minimise souvent des fonctions coûts de toute sorte. Pour résoudre (A.5.1) il existe d'autres algorithmes que le primal-dual lui ressemblant. Citons par exemple l'algorithme ADMM ou l'algorithme Douglas-Rachford, qui sont tous plus ou moins performants selon le problème qu'ils résolvent. Nous avons choisi l'algorithme primal-dual pour résoudre le transport optimal puisque l'article [Papadakis et al., 2014] semble indiquer que celui-ci est légèrement meilleur.

L'algorithme primal-dual utilise des opérateurs proximaux, définis pour une fonction  $f$  par

$$\text{Prox}_f(x) = \arg \min_{x'} f(x') + \frac{1}{2} \|x' - x\|^2$$

Ces opérateurs proximaux sont utilisés dans des algorithmes de minimisation de  $f$ . En effet, l'ajout du terme  $\|x' - x\|^2$  permet de bien convexifier la fonction  $f$  à minimiser, et d'avoir un algorithme parfois plus simple et qui converge.



On rappelle qu'on a défini le problème primal-dual (PD) comme étant le problème au point-selle

$$\inf_U \sup_V \mathcal{L}(U, V) \quad \text{avec} \quad \mathcal{L}(U, V) = \langle IU, V \rangle + \iota_C(U) - \iota_K(V)$$

On cherche une suite  $(U^n, V^n)$  telle que la valeur de  $\mathcal{L}$  croît à chaque itération de  $V$ , et décroît à chaque itération de  $U$ , jusqu'à converger au point selle. L'utilisation des opérateurs proximaux est intéressante pour ceci. L'algorithme primal-dual s'écrit tel quel

$$\begin{aligned} V^{n+1} &= \text{Prox}_{\sigma\iota_K}(V^n + \sigma IW^n) \\ U^{n+1} &= \text{Prox}_{\tau\iota_C}(U^n - \tau I^* V^{n+1}) \\ W^{n+1} &= U^{n+1} + \theta(U^{n+1} - U^n) \end{aligned}$$

avec  $\sigma$  et  $\tau$  deux constantes positives qui représentent les pas d'incrément. Cet algorithme se simplifie puisque l'opérateur proximal d'une fonction caractéristique d'un ensemble convexe est la projection sur cet ensemble. L'algorithme se réécrit alors

**Algorithme A.5.1** (*Algorithme primal-dual pour le transport optimal*)

Tant qu'on n'a pas atteint le critère de convergence, faire :

1.  $V^{n+1} = P_K(V^n + \sigma IW^n)$ .
2.  $U^{n+1} = P_C(U^n - \tau I^* V^{n+1})$
3.  $W^{n+1} = U^{n+1} + \theta(U^{n+1} - U^n)$

La distance de Wasserstein est donnée par la valeur au point-selle,  $\mathcal{W}_2(\rho_0, \rho_1)^2 = \langle U^n, V^n \rangle$ .

On détaille dans la suite les calculs des opérateurs de projection  $P_C$  et  $P_K$ .

L'ajout de la variable  $W$  permet d'avoir un algorithme qui converge mieux. Prendre  $\theta = 1$  est ce qu'il y a de plus efficace. De plus, les valeurs de  $\tau$  et  $\sigma$  sont contraints par  $\tau\sigma\|I\|^2 \leq 1$ , sinon l'algorithme risque de ne pas converger. Quelques tests seront nécessaires pour savoir quelles valeurs  $\sigma, \tau$  utiliser. C'est ce que l'on fera ensuite. Un critère d'arrêt est également à choisir. On comparera aussi cet algorithme avec l'algorithme du gradient conjugué. Détaillons d'abord les trois étapes de l'algorithme Primal-Dual (A.5.1).

**Étape 1 : Projection orthogonale sur  $K$ .** La première étape de l'algorithme consiste à projeter  $V^n + \sigma IW^n$  sur  $K$ . On rappelle que  $K$  est le paraboloïde défini par

$$K = \left\{ V = (a, b) : a + \frac{|b|^2}{2} \leq 0 \right\}$$

**Proposition A.5.2** (*Projection sur  $K$* )

La projection de  $\tilde{V} = (\tilde{a}, \tilde{b}) \notin K$  sur  $K$  est  $P_K(\tilde{V}) = (a, b)$ , avec la norme  $|b|$  de  $b$  la seule solution positive de l'équation

$$X^3 + 2(1 + \tilde{a})X - 2|\tilde{b}| = 0 \tag{A.5.2}$$

$a$  est donné par  $a = -|b|^2/2$  et enfin  $b$  est donné par  $b = \tilde{b}/(1 + \tilde{a} - a)$ .

**Démonstration :** Soit  $\tilde{V} = (\tilde{a}, \tilde{b}) \notin K$ . On cherche  $V = P_K(\tilde{V})$  la projection de  $V$  sur  $K$ , qui existe et est unique puisque  $K$  est convexe. Le vecteur  $\tilde{V} - P_K(\tilde{V})$  est parallèle au vecteur  $(1, b)$  normal à  $K$  en  $V$ , donc

$$\exists \lambda > 0: \quad \begin{pmatrix} \tilde{a} - a \\ \tilde{b} - b \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ b \end{pmatrix}$$

ce qui donne que  $\tilde{b} = (1 + \tilde{a} - a)b$ . Comme  $a = -b^2/2$ , on peut facilement trouver la norme de  $b$  en résolvant l'équation (A.5.2). S'en suit la valeur de  $a$  et puis celle de  $b$ . ■

**Étape 2 : Projection sur  $C$ .** La deuxième consiste à projeter  $U^n - \tau I^* V^{n+1}$  sur  $C$ ,  $C$  étant défini par

$$C = \left\{ U = (\rho, m) \text{ tel que } \begin{array}{l} \operatorname{div}_{t,x}(\rho, m) = 0, \\ \rho(0) = \rho_0, \quad \rho(1) = \rho_1, \\ m(x=0) = m(x=1) = 0 \end{array} \right\}.$$

### Proposition A.5.3

La projection de  $U \notin C$  sur  $C$  est donnée par

$$P_C(U) = U - \nabla \Phi$$

avec  $\Phi$  résolvant le Laplacien aux conditions aux bords de Neumann suivant :

$$\begin{cases} \Delta \Phi = \operatorname{div}(U) \\ \nabla \Phi \cdot \vec{n} = (\rho(0) - \rho_0, \rho(1) - \rho_1, m(x=0), m(x=1)) \end{cases}$$

La condition aux bords  $\nabla \Phi \cdot \vec{n}$  est exprimée sur  $\{t=0\} \times \{t=1\} \times \{x=0\} \times \{x=1\}$ .

**Démonstration :** La preuve est similaire à celle de la proposition A.3.2 donnant la projection sur  $TC$ . ■

**Étape 3 : mise à jour de  $W$ .** La variable  $W$  est calculée, pour  $\theta = 1$ , par  $W^{n+1} = 2U^{n+1} - U^n$ .

**Résumé de l'algorithme Primal-Dual.** L'algorithme primal-dual pour la résolution du problème de transport optimal est un algorithme itératif, chaque itération demandant une inversion d'un Laplacien spatio-temporel, ce qui est assez coûteux. La projection sur  $K$  s'avère être beaucoup plus rapide, notamment puisque tous les points n'ont pas à être projetés dans  $K$ .

Il reste à définir les valeurs des constantes  $\sigma$  et  $\tau$  ainsi qu'un critère d'arrêt. Ensuite, on comparera l'algorithme avec la descente du gradient.

**Choix des constantes  $\sigma$  et  $\tau$ .** Le choix des paramètres  $\sigma$  et  $\tau$  jouent un grand rôle dans la vitesse de convergence de l'algorithme. La théorie indique juste que si  $\sigma\tau\|I\|^2 \leq 1$ , l'algorithme converge (en fait, dans le cas où les fonctions sont fortement convexes, voir [Chambolle et Pock, 2011]). On constate ainsi qu'il faut choisir  $\sigma$  et  $\tau$  de telle sorte que

$$\sigma\tau = \frac{1}{\|I\|^2}.$$

On peut calculer  $\|I\|^2 = \max\{\operatorname{Sp}(I^*I)\}$  ( $\operatorname{Sp}$  désigne le spectre) :

**Proposition A.5.4**

$$I^*IU = U + \left( \frac{dt^2}{4} \partial_{tt}^2 \rho \right) + \left( \frac{dx^2}{4} \Delta m \right)$$

Et  $\max\{Sp(I^*I)\} = 1$ .

**Démonstration :** Sans donner explicitement  $I^*I$ , rappelons que  $I$  interpole  $\mathcal{U}$  sur  $\mathcal{V}$  et que  $I^*$  interpole  $\mathcal{V}$  sur  $\mathcal{U}$ . On se retrouve avec une double interpolation ;  $I^*I$  est en fait l'identité plus la discrétisation d'un Laplacien-Neumann. Étant donné que les valeurs propres du Laplacien sont strictement négatives, sauf pour les fonctions constantes, la plus grande valeur propre de  $I^*I$  vaut 1. ■

Les paramètres  $\sigma$  et  $\tau$  sont choisis comme dans [Papadakis et al., 2014]. L'algorithme est lancé avec 3000 itérations, quel que soit le jeu de paramètre il devrait atteindre l'optimum  $(U^*, V^*)$ . Ensuite, les paramètres sont choisis pour qu'au bout de 500 itérations, la valeur de  $\|U^n - U^*\|^2 + \|V^n - V^*\|^2$  soit minimale.

Ces tests donnent environ  $\sigma = 0,03$  et par conséquent  $\tau = 1/\sigma$ , cf. figure A.3 où l'on trace la valeur de  $\|U^n - U^*\|^2 + \|V^n - V^*\|^2$  pour  $n = 500$ . Nous avons utilisé les mêmes  $\rho_0$  et  $\rho_1$  que ceux donnés en figure A.2. Nous avons fait le même test pour des entrées et sorties différentes, et on retrouve sensiblement le même résultat.

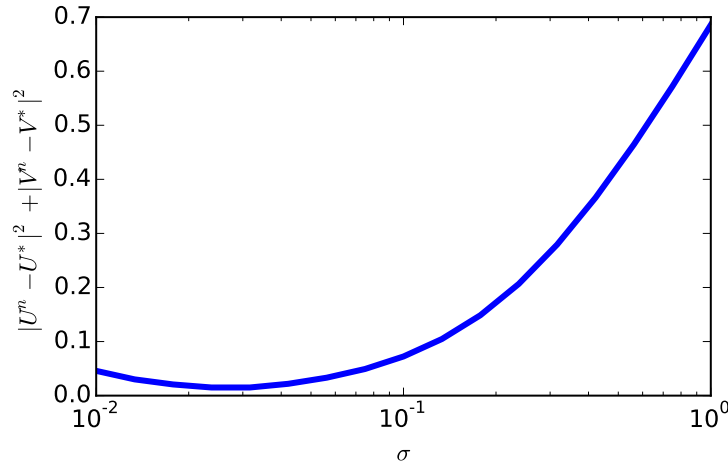


FIGURE A.3 – Valeur de  $\|U^n - U^*\|^2 + \|V^n - V^*\|^2$  en fonction de la valeur de  $\sigma$  utilisé dans l'algorithme primal dual, pour  $n = 500$ . On prend  $\tau = 1/\sigma$ .

**Le gradient.** Le gradient numérique de la fonction  $\mathcal{W}_2$  est approximatif : on calcule  $\Phi$  tel que  $I^*V = -\nabla\Phi$  et le gradient de la distance de Wasserstein par rapport à  $\rho_0$  est  $\Phi$  à  $t = 0$ .

**Critère d'arrêt** Il n'y a pas de critères d'arrêt précis. Dans l'article de Benamou et Brenier, les auteurs proposent d'utiliser la valeur de

$$\frac{\iint_{[0,1] \times \Omega} \rho [\partial_t \Phi + |\nabla \Phi|^2 / 2]}{\iint_{[0,1] \times \Omega} \rho |\nabla \Phi|^2}$$

avec  $\Phi$  tel que  $\nabla \Phi = -V$ . Ce critère d'arrêt mesure bien une certaine proximité à la solution, mais nous en utiliserons d'autres basées sur des heuristiques. En effet, ce critère doit valoir 0 à l'optimum

continu, mais pas dans l'optimum discret.

Étant donné qu'on réalise des approximations, nous avons deux quantités à contrôler : l'erreur d'approximation sur la valeur de la distance et l'erreur d'approximation sur la valeur du gradient. La valeur de la distance à l'itération  $n$  s'écrira  $L^n := \mathcal{L}(U^n, V^n)$ , et on notera la valeur du gradient par  $G^n = \nabla \Phi^n$ . On souhaiterait donc contrôler  $|L^n - L^*|$  et  $\|G^n - G^*\|$ , l'étoile désignant la valeur au point-selle, valeur que l'on peut approcher très précisément en faisant  $10^6$  itérations. En pratique pour calculer un transport optimal nous n'aurons pas la valeur de ce point-selle. Nous nous référons donc à la différence entre deux itérations. Pour estimer le lien entre la différence entre deux itérations et la distance au point-selle, nous traçons sur la figure A.4 on trace  $|L^{n+1} - L^n|$ ,  $|L^n - L^*|$ ,  $\|G^{n+1} - G^n\|$ ,  $\|G^n - G^*\|$ . On remarque que

$$|L^n - L^*| \leq 2.10^5 \times |L^{n+1} - L^n|$$

$$\|G^n - G^*\| \leq 10^2 \times \|G^{n+1} - G^n\|.$$

Ces constantes sont purement heuristiques et on n'est pas à l'abri de se tromper, mais cela semble être une bonne approximation. Ainsi, en fonction de la précision  $\epsilon$  que l'on souhaite, on arrêtera les algorithmes en fonction de la différence entre deux itérations. L'algorithme s'arrête lorsque

$$2.10^5 \times |L^{n+1} - L^n| < \epsilon \text{ et } 10^2 \times \|G^{n+1} - G^n\| < \epsilon.$$

**Comparaison avec l'algorithme de descente du gradient** On compare les algorithmes Primal-Dual et de descente du gradient. À convergence, les deux algorithmes doivent retrouver la même chose. Cependant, ils ne convergent pas du tout à la même vitesse, comme peut l'attester la figure A.5. Cette figure montre comment évolue  $\|U^n - U\|$ , en fonction de l'itération  $n$ . Il faut beaucoup moins d'itération pour l'algorithme Primal-Dual que pour l'algorithme de gradient conjugué, pour une erreur donnée.

## A.6 Régularisation d'entropie

Les algorithmes précédemment utilisés peuvent être longs à calculer. Des relaxations de la distance de Wasserstein existent et permettent un calcul plus rapide. On pense notamment à la relaxation d'entropie, proposée dans [Cuturi, 2013]. Cette formulation se base sur la version de Kantorovich (DK) du problème optimal, en ajoutant l'entropie de la mesure, c'est-à-dire que le problème à résoudre est

$$\min_{\gamma \in \Pi(\rho_0, \rho_1)} K_\epsilon(\gamma) := \iint_{\Omega \times \Omega} \frac{|x - y|^2}{2} \gamma(x, y) \, dx dy + \epsilon \iint_{\Omega \times \Omega} \gamma \log(\gamma) \, dx dy$$

avec  $\Pi(\rho_0, \rho_1)$  les mesures de marginale  $\rho_0$  sur la première variable, et  $\rho_1$  sur la deuxième,

$$\Pi(\rho_0, \rho_1) = \left\{ \begin{array}{l} \gamma \geq 0: \forall \Phi \in C^\infty(\Omega), \iint_{\Omega \times \Omega} \Phi(x) \, d\gamma(x, y) = \int_{\Omega} \Phi(x) \rho_0(x) \, dx, \\ \iint_{\Omega \times \Omega} \Phi(y) \, d\gamma(x, y) = \int_{\Omega} \Phi(y) \rho_1(y) \, dy \end{array} \right\}.$$

L'ajout dans  $K_\epsilon$  du terme strictement convexe d'entropie  $\gamma \log(\gamma)$  stabilise le problème, notamment cela fait apparaître une procédure qui rend le problème plus simple à résoudre. Lorsque  $\epsilon \rightarrow 0$ , la

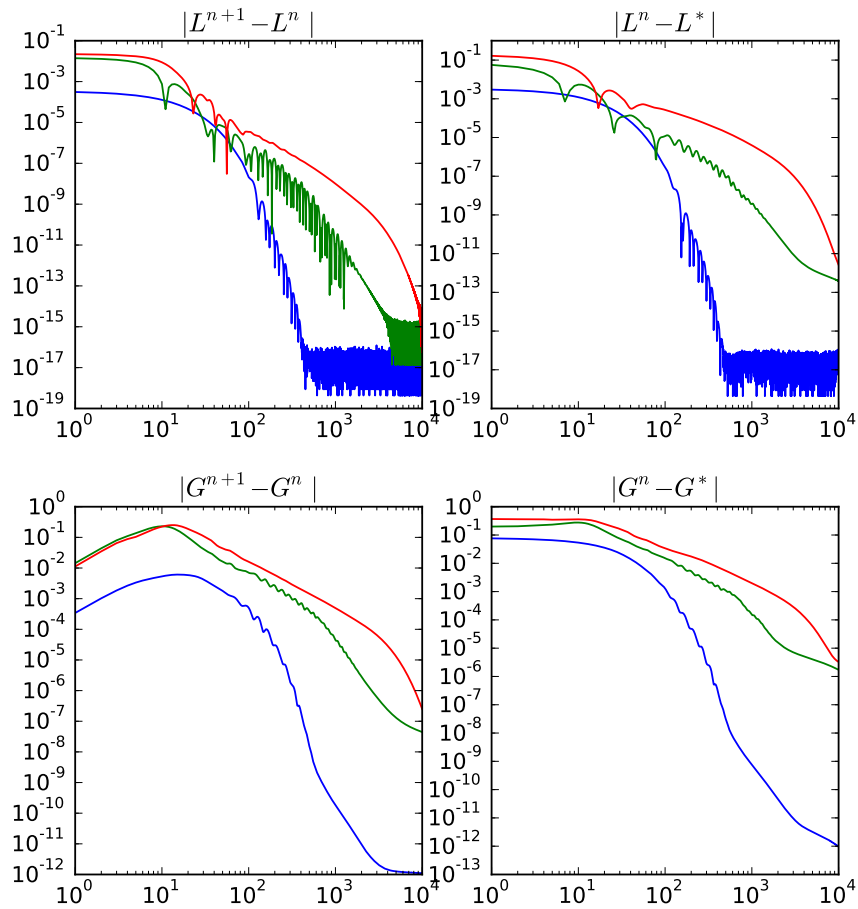


FIGURE A.4 – Tracé des courbes d'évolution de, en haut à gauche  $|L^{n+1} - L^n|$ , en haut à droite  $|L^n - L^*|$ , en bas à gauche  $\|G^{n+1} - G^n\|$  et en bas à droite  $\|G^n - G^*\|$ , en fonction des itérations, pour trois choix de  $\rho_0$  et  $\rho_1$ . Les  $L^*$  et  $G^*$  sont calculés en faisant  $10^6$  itérations.

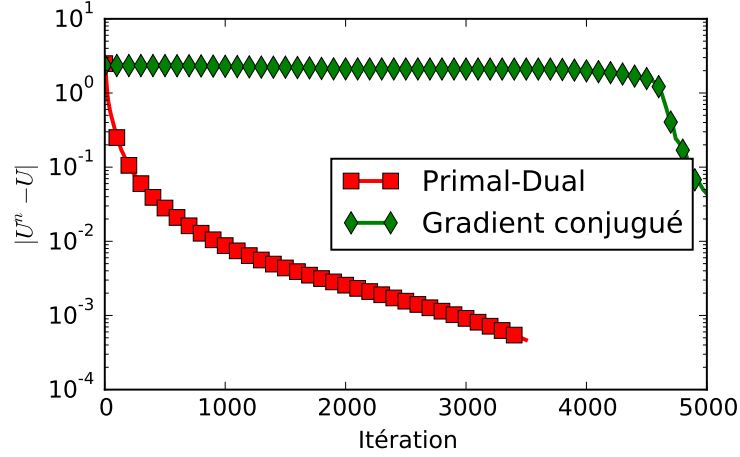


FIGURE A.5 – Comparaison de la vitesse de convergence de  $U^n$  vers  $U^*$ , c'est-à-dire qu'on trace  $\|U^n - U^*\|$  en fonction de l'itération  $n$ , pour l'algorithme de descente du gradient conjugué, voir la Section A.3, et pour l'algorithme Primal-Dual.

solution tend (au sens de la  $\Gamma$ -convergence [Carlier et al., 2015]) vers la solution du problème de Kantorovich, mais le temps de calcul de la procédure semble aussi tendre vers l'infini.

Avec les multiplicateurs de Lagrange traduisant les conditions sur  $\gamma$ , on résout le nouveau problème

$$\inf_{\gamma \geq 0} \sup_{\alpha(x), \beta(y)} K_\epsilon(\gamma) + \iint_{\Omega^2} [\alpha(x)(\gamma(x, y) - \rho_0(x)) + \beta(y)(\gamma(x, y) - \rho_1(y))] dx dy.$$

Une première équation d'Euler-Lagrange du point-selle est obtenue en différentiant l'équation ci-dessous par rapport à  $\gamma$  et est

$$\epsilon \log \left( \gamma(x, y) \exp \left( \frac{|x - y|^2}{2\epsilon} \right) \right) + \epsilon + \alpha(x) + \beta(y) = 0$$

En explicitant  $\gamma(x, y)$ ,

$$\gamma(x, y) = \underbrace{\exp \left( -\frac{c(x, y)}{\epsilon} \right)}_{M(x, y)} \underbrace{\exp \left( -\frac{1}{2} - \frac{\alpha(x)}{\epsilon} \right)}_{A(x)} \underbrace{\exp \left( -\frac{1}{2} - \frac{\beta(y)}{\epsilon} \right)}_{B(y)}$$

qui est automatiquement positif. Les  $A(x)$  et  $B(y)$  sont fixés par la condition de  $\gamma$  sur ses marginales. Ainsi, en intégrant sur  $\Omega$  cette équation une fois sur  $x$  et une autre sur  $y$ , on trouve

$$A(x) \int_{\Omega} M(x, y) B(y) dy = \rho_0(x), \quad B(y) \int_{\Omega} M(x, y) A(x) dx = \rho_1(y).$$

Numériquement, pour obtenir  $A(x)$  et  $B(y)$ , on utilise un algorithme de point fixe, qui donne

$$\begin{aligned} A^{n+1}(x) &= \frac{\rho_0(x)}{\int_{\Omega} M(x, y) B^n(y) dy} \\ B^{n+1}(y) &= \frac{\rho_1(y)}{\int_{\Omega} M(x, y) A^{n+1}(x) dx}. \end{aligned} \tag{A.6.1}$$

L'algorithme dépend de  $\epsilon$ . Lorsque  $\epsilon$  tend vers 0, il y a  $\Gamma$ -convergence du  $\gamma_\epsilon$  obtenu vers le  $\gamma$  correspondant au transport optimal. Cependant, numériquement, lorsque  $\epsilon$  tend vers 0, on remarque que l'algorithme met de plus en plus de temps à converger. De plus, des instabilités numériques apparaissent, en effet  $M(x, y)$  a des valeurs extrêmement petites ce qui engendre des valeurs de  $1/\int_{\Omega} M(x, y)B^n(y) dy$  trop élevées. Si bien que l'algorithme ne converge plus.

Pour un  $\epsilon$  de l'ordre de  $10^{-2}$ , l'algorithme converge très rapidement, en moins d'une seconde. C'est beaucoup plus rapide que l'algorithme Primal-Dual.

Lorsque le  $\epsilon$  est trop grand, l'algorithme converge encore plus rapidement, mais l'approximation est plus forte aussi. On trace sur la figure A.6 le barycentre pour cette distance pour différentes valeurs de  $\epsilon$ , et on remarque que pour  $\epsilon$  grand, le barycentre est lissé.

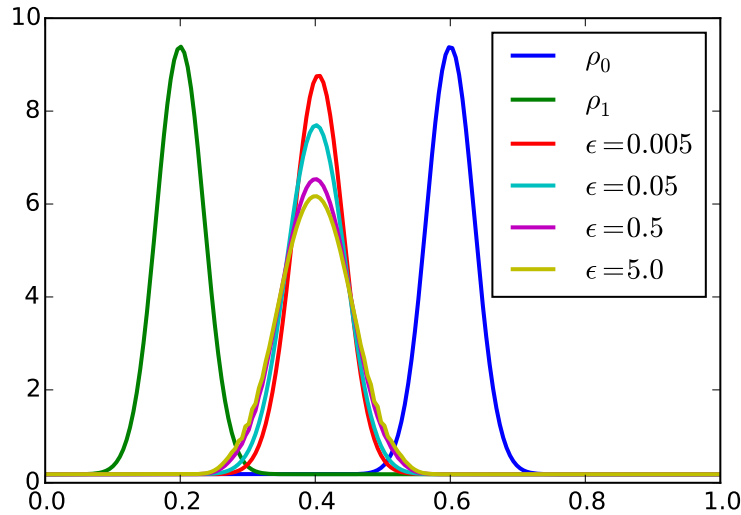


FIGURE A.6 – Comparaison pour différentes valeurs de  $\epsilon$  des barycentres entre les  $\rho_0$  et  $\rho_1$ .

Cela s'explique par l'équation du barycentre. La géodésique au sens de la distance régularisée est d'équation [Léonard, 2013] :

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho \nabla \Phi) = 0 \\ \partial_t \Phi + \frac{|\nabla \Phi|^2}{2} = \epsilon \Delta \Phi \end{cases}$$

La dernière équation est l'équation de Burgers potentielle, avec viscosité, qui se ramène à l'équation de la chaleur à l'aide du changement de variable d'Hopf-Cole. Ce tassement est en fait une diffusion qui peut ne pas être trop souhaitable car cela amène à de la perte d'information. C'est la contrepartie à un calcul très rapide de la distance de Wasserstein.

## A.7 Algorithme pour un calcul numérique de la distance WF.

La distance  $WF$  présentée plus tôt dans la partie 2.7 a été implémentée pour une utilisation numérique. Pour la calculer numériquement, on a utilisé le même algorithme Primal-Dual que celui utilisé pour la distance de Wasserstein classique présenté dans A.5. Si l'on effectue le changement de variable classique rendant les contraintes linéaires et la fonction coût convexe, c'est-à-dire  $m = \rho v$

et  $\xi = \rho g$ , on tombe sur

$$WF = \inf_{(\rho, m, \xi) \in C(\rho_0, \rho_1)} \iint \left[ \frac{|m|^2}{2\rho} + \delta \frac{\xi^2}{2\rho} \right] dt dx$$

avec

$$C = \{(\rho, m, \xi) : \partial_t \rho + \operatorname{div}(m) = \xi, \rho(0) = \rho_0, \rho(1) = \rho_1, m = 0 \text{ sur } \partial\Omega\}.$$

On note ainsi  $U = (\rho, m, \xi)$ ,  $J(U) = \iint |m|^2/2\rho + \delta\xi^2/2\rho$ . La version primale est :

$$WF(\rho_0, \rho_1) = \inf_U J(U) + \iota_C(U)$$

Pour obtenir l'algorithme primal-dual, on va effectuer les mêmes opérations que pour la distance de Wasserstein. Ainsi,

$$WF_\delta(\rho_0, \rho_1) = \inf_U \sup_V \langle U, V \rangle - J^*(V) + \iota_C(U)$$

On pose le produit scalaire entre deux éléments  $V, \tilde{V}$  par

$$\langle V, \tilde{V} \rangle = a\tilde{a} + b \cdot \tilde{b} + \delta c\tilde{c}.$$

Dans ce cas,

### Proposition A.7.1

$$J^*(V) = \iota_K(V) \text{ avec } K = \{(a, b, c) : a + \frac{|b|^2}{2} + \delta \frac{c^2}{2} \leq 0\}$$

**Démonstration :** On rappelle que

$$\begin{aligned} J^*(V) &= \sup_U \langle IU, V \rangle - \mathcal{J}(IU) \\ &= \sup_{(\rho, m, \xi)} a\rho + bm + \delta c\xi - \mathcal{J}(\rho, m, \xi). \end{aligned}$$

On partage le sup en trois cas : soit  $\rho > 0$ , soit  $(\rho, m, \xi) = (0, 0, 0)$ , soit le restant des cas. Dans le restant des cas,  $\mathcal{J}$  vaut  $+\infty$  et donc le sup sur cette partie vaut  $-\infty$ . Si  $(\rho, m, \xi) = (0, 0, 0)$  alors le sup vaut 0. Donc,

$$J^*(V) = \max \left( 0, \sup_{\rho > 0} a\rho + bm + \delta c\xi - \frac{|m|^2}{2\rho} + \delta \frac{\xi^2}{2\rho} \right).$$

et alors, en faisant ressortir le point critique,

$$a\rho + bm + \delta c\xi - \frac{|m|^2}{2\rho} + \delta \frac{\xi^2}{2\rho} = \rho \left[ a + \frac{b^2}{2} + \delta \frac{c^2}{2} - \frac{1}{2} \left( \left( b - \frac{m}{\rho} \right)^2 + \delta \left( c - \frac{\xi}{\rho} \right)^2 \right) \right].$$

Ainsi, si  $a + \frac{b^2}{2} + \delta \frac{c^2}{2} \leq 0$  alors  $J^*(V) = 0$ , et sinon  $J^*(V) = +\infty$ . ■

Avant de définir l'algorithme, on rappelle que les variables vivent sur des grilles différentes. Ainsi,  $U = (\rho, m, \xi)$  vit sur  $\mathcal{G}_s^t \times \mathcal{G}_s^x \times \mathcal{G}_c$  avec  $\mathcal{G}_s^t$  et  $\mathcal{G}_s^x$  les mêmes grilles décalées qu'avant, et  $\mathcal{G}_c$  une grille centrée. De même,  $V = (a, b, c)$  vit sur  $(\mathcal{G}_c)^3$ . Enfin, on a toujours l'opérateur d'interpolation  $I$  entre les deux variables, d'adjoint  $I^*$ . L'algorithme primal-dual s'écrit sans surprise, puisque les notations sont les mêmes qu'avant,



**Algorithme A.7.2 (Algorithme primal-dual pour WF)**

Tant qu'on n'a pas atteint le critère de convergence, faire :

1.  $V^{n+1} = P_K(V^n + \sigma IW^n)$ .
2.  $U^{n+1} = P_C(U^n - \tau I^* V^{n+1})$
3.  $W^{n+1} = U^{n+1} + \theta(U^{n+1} - U^n)$

La distance est donnée par la valeur au point-selle,  $WF(\rho_0, \rho_1)^2 = \langle IU^n, V^n \rangle$ .

La projection sur  $K$  est définie par :

**Proposition A.7.3**

Soit  $\tilde{V} = (\tilde{a}, \tilde{b}, \tilde{c}) \notin K$ . Alors la projection de  $\tilde{V}$  sur  $K$  (sur le bord de  $K$ ) est défini ainsi :

1. Soit  $d = \sqrt{|b|^2 + \delta c^2}$ ,  $\tilde{d} = \sqrt{|\tilde{b}|^2 + \delta \tilde{c}^2}$ ,
2.  $d$  est la solution positive de

$$X^3 + 2(1 + \tilde{a})X - 2\tilde{d} = 0 \quad (\text{A.7.1})$$

3.  $a$  vaut  $-d^2/2$ ,  $\lambda$  vaut  $\tilde{a} - a$ ,  $b$  vaut  $\tilde{b}/(1 + \lambda)$  et  $c = \tilde{c}/(1 + \lambda)$ .

**Démonstration :** La projection, définie par

$$\inf_{V \in K} \|V - \tilde{V}\|^2$$

peut être recalculé par un multiplicateur de Lagrange, sur la condition que  $a + b^2/2 + \delta c^2/2 = 0$ ,

$$\inf_V \sup_\lambda \frac{1}{2} \|V - \tilde{V}\|^2 + \lambda (a + b^2/2 + \delta c^2/2).$$

Le point-selle du Lagrangien vérifie les conditions d'Euler-Lagrange,

$$\begin{cases} \tilde{a} - a &= \lambda \\ \tilde{b} &= (1 + \lambda)b \\ \tilde{c} &= (1 + \lambda)c \end{cases}$$

On pose  $d = \sqrt{|b|^2 + \delta c^2}$ ,  $\tilde{d} = \sqrt{|\tilde{b}|^2 + \delta \tilde{c}^2}$ , et alors  $d$  est la seule solution positive de (A.7.1), et les valeurs de  $a, b$  et  $c$  suivent. ■

La projection sur  $C$  est donnée par la proposition suivante :

**Proposition A.7.4 (Projection sur C)**

La projection de  $U \notin C$  sur  $C$  est

$$P_C(U) = U - \begin{pmatrix} \partial_t \Phi \\ \partial_x \Phi \\ \Phi/\delta \end{pmatrix}$$

avec  $\Phi$  résolvant le problème suivant

$$\begin{cases} \Delta_{t,x} \Phi - \Phi/\delta = \text{div}_{t,x}(\rho, m) - \xi \\ \nabla \Phi \cdot \vec{n} = (\rho(0) - \rho_0, \rho(1) - \rho_1, m(0), m(1)) \end{cases}$$

**Démonstration :** Tout comme les deux propositions précédentes A.3.2 et A.5.3, le vecteur  $U - P_C(U)$  appartient à  $\ker(F)^\perp$  avec

$$F(U) = \operatorname{div}_{t,x}(\rho, m) - \xi$$

et donc  $U - P_C(U) \in \operatorname{Im}(F^*)$ , d'où la proposition. ■

À noter que dans l'article original [Chizat et al., 2015b], les auteurs ont implémenté un algorithme qui repose sur le schéma Douglas-Rachford, et non sur le schéma Primal-Dual. On a implémenté le schéma Douglas-Rachford car on l'avait déjà fait pour l'algorithme du calcul de la distance de Wasserstein. Le schéma Douglas-Rachford est légèrement différent et les performances équivalentes voire un peu meilleures pour l'algorithme Primal-Dual, comme le montre [Papadakis et al., 2014].

De plus, il est récemment sorti une régularisation entropique prometteuse de la distance WF [Chizat et al., 2016], mais que nous n'avons pas implémentée.

**Valeurs de  $\sigma$  et  $\tau$**  De même qu'avant, une condition suffisante de convergence est  $\tau\sigma\|I\| \leq 1$ . On montre comme avant que  $\|I\| = 1$ . Pour trouver les valeurs de  $\tau$  et  $\sigma$ , on laisse tourner l'algorithme un nombre suffisant de fois pour atteindre la convergence, avant de le relancer avec différentes valeurs de  $\sigma$  et  $\tau = 1/\sigma$  pour voir quelle valeur de  $\sigma$  donne la convergence la plus rapide.

De même qu'avant, on trouve que  $\sigma = 0.03$  est une bonne valeur.



# Annexe B

## La méthode de l'adjoint

### Contenu du chapitre

B.1	Gradient d'une fonction impliquant un modèle . . . . .	159
B.2	Le modèle tangent . . . . .	160
B.3	Adjoint continu . . . . .	161
B.4	Adjoint discret . . . . .	162
B.5	Équations adjointes pour modèle de Saint-Venant avec fond . . . . .	162

### B.1 Gradient d'une fonction impliquant un modèle

Les méthodes efficaces de minimisation d'une fonction coût  $\mathcal{J}$  impliquent le calcul du gradient. Ces fonctions  $\mathcal{J}$  mettent souvent en jeu un modèle qu'il faut donc pouvoir différentier et dont il faut calculer le gradient.

Le but de cette annexe est de montrer comment calculer le gradient de

$$\mathcal{J}(h_0) = \|\mathcal{M}(h_0) - h^{\text{obs}}\|_2^2$$

avec une seule observation  $h^{\text{obs}}$  prise au temps  $T$ , et  $\mathcal{M}(h_0)$  un modèle de condition initiale  $h_0$ , qui calcule la valeur de  $h(T)$ . On prendra le modèle de Saint-Venant avec vitesse initiale nulle,

$$\mathcal{M}: \begin{cases} \partial_t h + \text{div}(hu) = 0 \\ \partial_t u + u \cdot \nabla u + g \nabla h = 0 \\ h(0, x) = h_0(x), \quad u(0, x) = 0, \quad u \cdot \vec{n} = 0 \text{ sur } \partial\Omega. \end{cases} \quad (\text{B.1.1})$$

La différentielle de  $\mathcal{J}$  est définie pour une perturbation  $\eta$  par

$$D\mathcal{J}[h_0] \cdot \eta_0 = \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{J}(h_0 + \varepsilon \eta_0) - \mathcal{J}(h_0)}{\varepsilon}.$$

Cette forme linéaire, fonction de  $\eta_0$ , s'écrit par le théorème de représentation de Riesz par un gradient  $\text{grad}\mathcal{J}[h_0]$  dépendant de  $h_0$  si elle est continue (ce qui arrive lorsque le modèle  $\mathcal{M}$  est suffisamment régulier) :

$$D\mathcal{J}[h_0] \cdot \eta_0 = \langle \eta_0, \text{grad}\mathcal{J}[h_0] \rangle_2.$$

Il n'y a en fait pas de formule explicite du gradient à cause du modèle impliqué. Une première méthode pour approcher le gradient du modèle discrétisé est d'utiliser les différences finies. En effet, si l'on prend  $\eta_0$  valant 1 au point  $x_i$  et 0 ailleurs, on a la valeur du gradient au point  $x_i$  puisque

$$\langle \eta_0, \text{grad} \mathcal{J} \rangle = \text{grad} \mathcal{J}[h_0](x_i)$$

et donc il suffit de calculer

$$\frac{\mathcal{J}(h_0 + \varepsilon \eta) - \mathcal{J}(h_0)}{\varepsilon} \quad (\text{B.1.2})$$

qui tend pour  $\varepsilon$  tendant vers 0 vers  $\text{grad} \mathcal{J}[h_0](x_i)$ . Cependant, cette méthode a deux défauts. D'abord, elle n'offre qu'une approximation du gradient puisque (B.1.2) sera calculé avec un  $\varepsilon$  non nul. Ensuite, cette méthode nécessite autant d'intégrations du modèle que de points du maillage, il s'avère que cela rend le calcul du gradient long.

Nous allons maintenant expliciter le gradient du modèle continu avant de parler du gradient du modèle discrétisé. Pour ceci, nous allons d'abord expliciter  $\mathcal{J}(h_0 + \varepsilon \eta_0)$  à l'aide du modèle tangent.

## B.2 Le modèle tangent

Pour différentier  $\mathcal{J}(h_0 + \varepsilon \eta_0)$ , nous avons besoin de différentier  $\mathcal{M}(h_0 + \varepsilon \eta_0)$ . Pour ceci, on utilise le modèle tangent, en  $h_0$ , et que l'on écrit  $\frac{\partial \mathcal{M}}{\partial h}[h_0].\eta_0$  ou  $\mathbf{M}[h_0].\eta_0$ , et on a

$$\mathcal{M}(h_0 + \varepsilon \eta_0) = \mathcal{M}(h_0) + \varepsilon \mathbf{M}[h_0].\eta_0 + \mathcal{O}(\varepsilon).$$

Le modèle tangent calcule les perturbations de  $h(T)$  au temps  $T$  lorsqu'une perturbation  $\eta_0$  a été insufflée au temps initial. C'est une linéarisation de  $\mathcal{M}$  autour de  $h_0$ .

Concrètement, son calcul se fait par linéarisation des variables impliquées. On dira que les variables  $h_\varepsilon(t, x)$  et  $u_\varepsilon(t, x)$  solutions de  $\mathcal{M}(h_0 + \varepsilon \eta_0)$  se linéarisent à l'aides des solutions  $h(t, x)$  et  $u(t, x)$  de  $\mathcal{M}(h_0)$  par

$$\begin{aligned} h_\varepsilon(t, x) &= h(t, x) + \varepsilon \eta(t, x) + \mathcal{O}(\varepsilon) \\ u_\varepsilon(t, x) &= u(t, x) + \varepsilon v(t, x) + \mathcal{O}(\varepsilon) \end{aligned}$$

En regardant à  $t = 0$  nous trouvons les conditions initiales qui sont  $\eta(0, x) = \eta_0(x)$  et  $v(0, x) = 0$ . En injectant ces variables dans les équations de Saint-Venant (B.1.1), nous avons au premier ordre les équations que vérifient  $\eta(t, x)$  et  $v(t, x)$ , ce sont les équations du modèle tangent  $\mathbf{M}[h_0].\eta_0$  :

$$\mathbf{M}[h_0].\eta_0: \begin{cases} \partial_t \eta + \text{div}(\eta u) + \text{div}(h v) = 0 \\ \partial_t v + \nabla(u \cdot v) + g \nabla \eta = 0 \\ \eta(0, x) = \eta_0(x), \quad v(0, x) = 0, \quad v \cdot \vec{n} = 0 \text{ sur } \partial \Omega. \end{cases} \quad (\text{B.2.1})$$

Nous avons donc que

$$\begin{aligned} \mathcal{J}(h_0 + \varepsilon \eta_0) &= \|\mathcal{M}(h_0 + \varepsilon \eta_0) - h^{\text{obs}}\|_2^2 \\ &= \|\mathcal{M}(h_0) + \varepsilon \mathbf{M}[h_0].\eta_0 + \mathcal{O}(\varepsilon) - h^{\text{obs}}\|_2^2 \\ &= \|\mathcal{M}(h_0) - h^{\text{obs}}\|_2^2 + 2\varepsilon \langle \mathbf{M}[h_0].\eta_0, \mathcal{M}(h_0) - h^{\text{obs}} \rangle_2 + \mathcal{O}(\varepsilon). \end{aligned}$$

Ceci nous donne la différentielle,

$$D\mathcal{J}(h_0).\eta_0 = 2 \langle \mathbf{M}[h_0].\eta_0, \mathcal{M}(h_0) - h^{\text{obs}} \rangle_2,$$

mais pas le gradient, puisque pour avoir le gradient il faudrait que  $\eta_0$  soit isolé. Pour isoler  $\eta_0$ , nous utilisons la méthode de l'adjoint.

### B.3 Adjoint continu

Pour calculer le gradient, on utilise la technique de l'adjoint, développée par J.L.Lions pour la théorie du contrôle optimal [Lions, 1971] et introduite par [Le Dimet et Talagrand, 1986] pour l'assimilation de données. C'est une technique qui se résume à un modèle décrit par un système d'équations différentielles rétrogrades. Si l'on note par  $\mathbf{M}^*[h_0]$  le modèle adjoint, celui-ci se définit par rapport au modèle tangent par :

$$\forall \eta_0, d, \quad \langle \mathbf{M}[h_0].\eta_0, d \rangle_2 = \langle \eta_0, \mathbf{M}^*[h_0].d \rangle_2.$$

#### Théorème B.3.1 (Modèle adjoint de Saint-Venant)

Soit  $h(t, x)$  et  $u(t, x)$  solution du modèle direct  $\mathcal{M}(h_0)$ . Soit  $p(t, x)$  et  $q(t, x)$  solutions du modèle adjoint

$$\mathbf{M}^*[h_0].d: \quad \begin{cases} \partial_t p + u \cdot \nabla p + g \operatorname{div}(q) = 0 \\ \partial_t q + \operatorname{div}(q)u + h \nabla p = 0 \\ p(T, x) = d(x), \quad q(T, x) = 0, \quad q(t) \cdot \vec{n} = 0 \text{ sur } \partial\Omega. \end{cases} \quad (\text{B.3.1})$$

Alors la valeur de  $p$  à  $t = 0$  est la valeur de  $\mathbf{M}^*[h_0].d$ , dans le sens qu'on a

$$\langle \mathbf{M}[h_0].\eta_0, d \rangle_2 = \langle \eta_0, p(0) \rangle_2. \quad (\text{B.3.2})$$

**Démonstration :** On dénote par  $p(t, x)$  et  $q(t, x)$  les variables adjointes respectivement à  $\eta$  et  $v$ . En faisant la somme du produit scalaire euclidien (sur  $[0, T] \times \Omega$ ) entre l'équation de  $\eta$  de (B.2.1) et  $p$ , et le produit scalaire entre l'équation de  $v$  et  $q$ , nous obtenons

$$\langle \partial_t \eta + \operatorname{div}(\eta u) + \operatorname{div}(h v), p \rangle_2 + \langle \partial_t v + \nabla(u \cdot v) + g \nabla \eta, q \rangle_2 = 0.$$

Puisque le produit scalaire est une intégration sur  $[0, T] \times \Omega$ , on peut effectuer les intégrations par parties en espace et en temps pour faire passer les dérivées sur  $p$  et  $q$ . En prenant en compte le fait que  $u$  et  $v$  sont tangents au bord, on a

$$\begin{aligned} - \langle \eta, \partial_t p \rangle_2 + \int_{\Omega} [p(T)\eta(T) - p(0)\eta(0)] - \langle \eta u + h v, \nabla p \rangle_2 + \oint_{\partial\Omega} g \eta q \cdot \vec{n} \, ds \\ - \langle v, \partial_t q \rangle_2 + \int_{\Omega} [q(T)v(T) - q(0) \cdot v(0)] \, dx - \langle u \cdot v + g \eta, \operatorname{div}(q) \rangle_2 = 0 \end{aligned} \quad (\text{B.3.3})$$

D'une part,  $v(0) = 0$ ,  $\eta(0) = \eta_0$  et  $\eta(T) = \mathbf{M}[h_0].\eta_0$ . D'autre part, on isole  $\eta$  et  $v$  pour avoir

$$\begin{aligned} \langle \eta, \underbrace{\partial_t p + u \cdot \nabla p + g \operatorname{div}(q)}_{(a)} \rangle_2 + \langle v, \underbrace{\partial_t q + \operatorname{div}(q)u + h \nabla p}_{(b)} \rangle_2 + \oint_{\partial\Omega} g \eta q \cdot \vec{n} \, ds \\ = \int_{\Omega} [p(T)\mathbf{M}[h_0].\eta_0 + q(T) \cdot v(T) - p(0)\eta_0]. \end{aligned} \quad (\text{B.3.4})$$

Ainsi, on remarque que si l'on prend  $p$  et  $q$  tels que (a) et (b) sont nuls, que  $p(T) = d$ ,  $q(T) = 0$ , et que  $q \cdot \vec{n} = 0$  sur  $\partial\Omega$ , c'est-à-dire si nous prenons  $p$  et  $q$  vérifiant (B.3.1), alors il ne reste que :

$$\int_{\Omega} d \cdot \mathbf{M}[h_0].\eta_0 = \int_{\Omega} p(0)\eta_0$$

qui se réécrit

$$\langle d, \mathbf{M}[h_0].\eta_0 \rangle_2 = \langle p(0), \eta_0 \rangle_2$$

On a donc prouvé que ce  $p(0)$  était égal à  $\mathbf{M}^*[h_0].d$ . ■

On peut remarquer que modèle adjoint (B.3.1) est une intégration rétrograde : on connaît la valeur des variables adjointes au temps  $T$  et on les calcule au temps  $t = 0$ . On remarque qu'il faut aussi connaître la valeur de  $h$  et  $u$  en tout temps pour résoudre l'équation puisqu'ils interviennent dans le modèle.

## B.4 Adjoint discret

En théorie, nous travaillons avec des données et équations discrétisées : on ne connaît et ne calcule les différentes données qu'en un nombre fini de points spatiaux et temporels. Le modèle  $\mathcal{M}$  a été discrétisé. Le modèle tangent découle directement du modèle direct discrétisé par linéarisation et est également discrétisé. Pour le modèle adjoint, on a deux solutions ;

- Soit on discrétise le modèle adjoint continu (B.3.1). Mais si la discrétisation est mal faite, alors la relation (B.3.2) définissant l'adjoint sera fautive à cause de l'erreur de discrétisation.
- Soit on écrit l'adjoint du modèle discret. Cette procédure a l'avantage de donner un gradient exact (à l'erreur machine près). L'idée est de reproduire la démarche présentée dans la preuve du théorème B.3.1 sur le modèle tangent discrétisé : il faut contrer les équations discrétisées par les variables adjointes, intégrer et effectuer les intégrations par partie (nommées "transformation d'Abel" en discret). C'est un processus long et fastidieux, même s'il existe quelques techniques pratiques qui rendent le travail plus efficace. Notons tout de même depuis 1990 l'apparition d'outils de différentiation automatique<sup>1</sup>, qui effectuent algorithmiquement cette tâche, notamment Tapenade [Hascœt et Pascual, 2004] qui effectue la différentiation automatique de codes écrits en Fortran 95.

## B.5 Équations adjointes pour modèle de Saint-Venant avec fond

Dans le chapitre 5 nous avons utilisé le modèle de Saint-Venant avec fond, fond décrit par la variable  $f$ . Les équations du modèle dépendent de  $h_0$ ,  $u_0$  et  $f$  et sont :

$$\mathcal{M}: \begin{cases} \partial_t h + \operatorname{div}(hu) = 0 \\ \partial_t u + u \cdot \nabla u + g \nabla(h + f) = 0 \\ h(0, x) = h_0(x), \quad u(0, x) = u_0(x), \quad u \cdot \vec{n} = 0 \text{ sur } \partial\Omega. \end{cases} \quad (\text{B.5.1})$$

Nous cherchons le gradient d'une fonction dépendant de  $h_0$ ,  $u_0$  et  $f$  définie par

$$\mathcal{J}(h_0, u_0, f) = \frac{1}{2} \|\mathcal{G}(h_0, u_0, f) - y^{\text{obs}}\|_2^2$$

avec  $\mathcal{G}$  l'observateur général d'observations qui envoie  $(h_0, u_0, f)$  sur l'altitude de la surface libre au temps  $T$ ,  $h(T) + f$ . Il faut pouvoir différentier  $\mathcal{G}$ , donnant le modèle tangent  $\mathbf{G}$ , et écrire son adjoint  $\mathbf{G}^*$ , de telle sorte que :

$$\begin{aligned} \mathcal{J}(h_0 + \varepsilon \eta_0, u_0 + \varepsilon v_0, f + \varepsilon m) &= \frac{1}{2} \|\mathcal{G}(h_0 + \varepsilon \eta_0, u_0 + \varepsilon v_0, f + \varepsilon m) - y^{\text{obs}}\|_2^2 \\ &= \frac{1}{2} \|\mathcal{G}(h_0, u_0, f) + \varepsilon \mathbf{G} \cdot (\eta_0, v_0, m) - y^{\text{obs}}\|_2^2 \\ &= \mathcal{J}(h_0, u_0, f) + \varepsilon \langle \mathbf{G} \cdot (\eta_0, v_0, m), d \rangle_2 + \mathcal{O}(\varepsilon) \\ &= \mathcal{J}(h_0, u_0, f) + \varepsilon \langle (\eta_0, v_0, m), \mathbf{G}^* d \rangle_2 + \mathcal{O}(\varepsilon) \end{aligned} \quad (\text{B.5.2})$$

1. Voir le site <http://www.autodiff.com> qui recense plusieurs de ces logiciels.

avec  $d = \mathcal{G}(h_0, u_0, f) - y^{\text{obs}}$ . Le  $\mathbf{G}^*d$  est donc le gradient  $\text{grad}\mathcal{J}$  de  $\mathcal{J}$ , qu'on récrit avec ses trois composantes

$$\text{grad}\mathcal{J} = \mathbf{G}^*d = \begin{pmatrix} \mathbf{G}_h^*d \\ \mathbf{G}_u^*d \\ \mathbf{G}_f^*d \end{pmatrix}$$

Nous allons maintenant expliciter  $\mathbf{G}^*$ , pour ceci nous aurons besoin du modèle tangent  $\mathbf{G}$ .

Le modèle tangent  $\mathbf{M}$  s'écrit pour une petite perturbation  $(\eta_0, v_0, m)$  de  $(h_0, u_0, f)$  par

$$\mathbf{M}[h_0, u_0, f].(\eta_0, v_0, m): \begin{cases} \partial_t \eta + \text{div}(\eta u) + \text{div}(h v) = 0 \\ \partial_t v + \nabla(u \cdot v) + g \nabla(\eta + m) = 0 \\ \eta(0, x) = \eta_0(x), \quad v(0, x) = v_0(x), \quad v \cdot \vec{n} = 0 \text{ sur } \partial\Omega. \end{cases} \quad (\text{B.5.3})$$

Alors,  $\mathbf{G}(\eta_0, v_0, m) = \eta(T) + m$ .

Ces équations ressemblent à celles du modèle tangent sans fond (B.2.1). En plus de la condition initiale  $v(0, x) = v_0(x)$ , l'équation en  $v$  a changé puisqu'un terme  $g \nabla m$  est apparu.

**Modèle adjoint.** Le modèle adjoint s'écrit alors :

#### Théorème B.5.1

*Nous posons les variables adjointes à  $\eta(t, x)$ ,  $v(t, x)$  et  $m(x)$  qui sont  $p(t, x)$ ,  $q(t, x)$  et  $r(x)$ . Les équations adjointes sont :*

$$\mathbf{M}^*[h_0, u_0, f].d: \begin{cases} \partial_t p + u \cdot \nabla p + g \text{div}(q) = 0 \\ \partial_t q + \text{div}(q)u + h \nabla p = 0 \\ r(x) = g \int_{[0, T]} \text{div}(q)(t, x) dt \\ p(T, x) = d(x), \quad q(T, x) = 0, \quad q(t) \cdot \vec{n} = 0 \text{ sur } \partial\Omega. \end{cases} \quad (\text{B.5.4})$$

Alors,

$$\text{grad}\mathcal{J}(h_0, u_0, f) = \mathbf{G}^*d = \begin{pmatrix} p(0) \\ q(0) \\ r + d \end{pmatrix}$$

**Démonstration :** De même que pour l'équation adjointe (B.3.1), nous controns la première équation de (B.5.3) avec  $p$ , la deuxième avec  $q$ , intégrons sur  $[0, T] \times \Omega$ , ajoutons, intégrons par partie, et isolons  $\eta$ ,  $v$  et  $m$ , ce qui donne :

$$\begin{aligned} & \langle \eta, \underbrace{\partial_t p + u \cdot \nabla p + g \text{div}(q)}_{(a)} \rangle_2 + \langle v, \underbrace{\partial_t q + \text{div}(q)u + h \nabla p}_{(b)} \rangle_2 + \oint_{\partial\Omega} g \eta q \cdot \vec{n} ds \\ & + \langle m, g \text{div}(q) \rangle_2 = \int_{\Omega} [p(T)\eta(T) + q(T) \cdot v(T) - p(0)\eta_0 - q(0)v_0]. \end{aligned} \quad (\text{B.5.5})$$

en ayant incorporé que  $\eta(0) = \eta_0$  et  $v(0) = v_0$ . Si nous prenons  $p$  tel que (a) est nul,  $p(T) = d$ ,  $q$  tel que (b) est nul,  $q(T) = 0$  et  $g \cdot \vec{n} = 0$  sur  $\partial\Omega$ , alors :

$$\int_{\Omega} \eta(T) d dx = \int_{\Omega} [p(0)\eta_0 + q(0)v_0 + \underbrace{m(x) \int_{[0, T]} g \text{div}(q) dt}_{r(x)}] dx$$



si bien que

$$\begin{aligned}
 \langle \mathbf{G}(\eta_0, v_0, m), d \rangle_2 &= \langle \eta(T) + m, d \rangle_2 \\
 &= \langle \eta(T), d \rangle_2 + \langle m, d \rangle_2 \\
 &= \int_{\Omega} [p(0)\eta_0 + q(0)v_0 + r m] dx + \langle m, d \rangle_2 \\
 &= \langle \eta_0, p(0) \rangle_2 + \langle v_0, q(0) \rangle_2 + \langle m, d + r \rangle_2
 \end{aligned}$$

et donc en reprenant les calculs de la différentiation de  $\mathcal{J}$ ,

$$\mathcal{J}(h_0 + \varepsilon\eta_0, u_0 + \varepsilon v_0, f + \varepsilon m) = \mathcal{J}(h_0, u_0, f) + \varepsilon [\langle \eta_0, p(0) \rangle_2 + \langle v_0, q(0) \rangle_2 + \langle m, d + r \rangle_2] + \mathcal{O}(\varepsilon)$$

et le gradient s'écrit

$$\text{grad}\mathcal{J} = \begin{pmatrix} p(0) \\ q(0) \\ r + d \end{pmatrix}. \quad \blacksquare$$

## Annexe C

# Dans le cas où les conditions aux bords ne sont pas de Neumann

Nous rappelons l'hypothèse dont nous avons eu besoin pour faire nos expériences. Soit  $\Phi$  le gradient  $\mathcal{L}^2$  en  $\rho_0$  d'une fonction  $\mathcal{J}$ , gradient défini par

$$\forall \eta, \mathcal{J}(\rho_0 + \varepsilon \eta) = \mathcal{J}(\rho_0) + \varepsilon \langle \eta, \Phi \rangle_2 + \mathcal{O}(\varepsilon)$$

Si  $\Phi$  vérifie la condition aux bords de Neumann homogène (on dira (CNH)), c'est-à-dire si

$$\rho \frac{\partial \Phi}{\partial \vec{n}} = 0 \text{ sur } \partial \Omega \quad (\text{CNH})$$

alors le gradient  $\mathcal{W}_2$  de  $\mathcal{J}(\rho_0)$  est

$$-\text{div}(\rho_0 \nabla \Phi).$$

Si  $\Phi$  n'est pas (CNH), le gradient  $\mathcal{W}_2$  n'existe généralement pas.

**Démonstration :** Supposons que le gradient  $\text{grad}_W \mathcal{J}$  existe. Il existe alors un  $\tilde{\Phi}$ , vérifiant (CNH), tel que

$$\text{grad}_W \mathcal{J} = -\text{div}(\rho_0 \nabla \tilde{\Phi})$$

Alors pour tout  $\eta = -\text{div}(\rho_0 \nabla \Psi)$  de  $T_{\rho_0} \mathcal{P}$ , ( $\Psi$  vérifie également (CNH)), nous avons

$$\begin{aligned} \langle \text{grad}_W \mathcal{J}, \eta \rangle_W &= \int_{\Omega} \rho_0 \nabla \tilde{\Phi} \cdot \nabla \Psi \\ &= - \int_{\Omega} \tilde{\Phi} \text{div}(\rho_0 \nabla \Psi) + \oint_{\partial \Omega} \underbrace{\Psi \rho_0 \frac{\partial \tilde{\Phi}}{\partial \vec{n}}}_{=0} \\ &= \int_{\Omega} \tilde{\Phi} \eta \\ &= \langle \tilde{\Phi}, \eta \rangle_2. \end{aligned}$$

Or par définition du gradient,  $\langle \text{grad}_W \mathcal{J}, \eta \rangle_W$  vaut la différentielle de  $\mathcal{J}$  en  $\rho_0$ , exprimé aussi par le gradient euclidien  $\Phi$  :

$$\langle \text{grad}_W \mathcal{J}, \eta \rangle_W = \langle \Phi, \eta \rangle_2$$

et comme ceci est valable pour tout  $\eta$ , nous avons  $\tilde{\Phi} = \Phi$ . Or  $\Phi$  et  $\tilde{\Phi}$  ne vérifient pas les mêmes conditions aux bords (le dernier vérifie (CNH) mais pas le premier) : c'est une absurdité et donc l'hypothèse initiale est fautive : le gradient de Wasserstein n'existe en fait pas. ■

Nous avons expliqué dans la remarque 3.3.9 que dans le cas d'une fonction coût écrite à l'aide de la distance de Wasserstein,  $\Phi$  était généralement la somme de potentiels de Kantorovich (de conditions aux bords de Neumann) éventuellement transformés par l'adjoint du modèle  $\mathbf{M}^*$ . Le problème est que cet adjoint ne préserve pas forcément cette propriété, c'est-à-dire que si nous avons  $\Phi^{\text{obs}}$  qui soit (CNH), il se peut que  $\mathbf{M}^*\Phi^{\text{obs}}$  ne soit pas (CNH). Alors, le gradient de Wasserstein n'existe plus.

Nous proposons la définition d'un autre produit scalaire pour lequel le gradient de Wasserstein existe parfois et vaut toujours  $-\text{div}(\rho_0 \nabla \Phi)$ .

## C.1 Définition du produit scalaire

L'espace tangent à  $\mathcal{P}(\Omega)$  en  $\rho_0$  était défini comme étant l'ensemble des éléments

$$\eta = -\text{div}(\rho_0 \nabla \Phi) \text{ avec } \frac{\partial \Phi}{\partial \vec{n}} = 0 \text{ sur } \partial \Omega.$$

Ce sont ces conditions aux bords de Neumann qui obligent le gradient euclidien à être également (CNH) pour qu'un gradient de Wasserstein existe. Dans le cas où le gradient euclidien  $\Phi$  a des conditions aux bords quelconques, nous changeons les conditions aux bords dans la définition de l'espace tangent pour trouver un gradient de Wasserstein. Ces conditions aux bords sont de Robin. Soit  $\alpha(x)$  défini sur  $\partial \Omega$  tel que le gradient euclidien  $\Phi$  vérifie sur  $\partial \Omega$  :

$$\rho_0 \left[ \frac{\partial \Phi}{\partial \vec{n}} + \alpha \Phi \right] = 0$$

et  $\alpha(x) = 0$  si  $\rho_0(x) \frac{\partial \Phi}{\partial \vec{n}}(x) = 0$  ou si  $\rho_0(x) \Phi(x) = 0$ . Alors, l'espace tangent est l'ensemble des éléments dont le potentiel vérifie ces conditions aux bords de Robin :

$$T_{\rho_0} \mathcal{P} = \{ \eta = -\text{div}(\rho \nabla \Psi), \quad \rho \frac{\partial \Psi}{\partial \vec{n}} + \rho \alpha \Psi = 0 \}.$$

Enfin, le produit scalaire est réécrit en prenant en compte ces conditions aux bords, comme on le fait dans la théorie d'inversion du Laplacien avec conditions aux bords de Robin :

$$\forall \eta = -\text{div}(\rho \nabla \Phi), \eta' = -\text{div}(\rho \nabla \Psi'), \langle \eta, \eta' \rangle = \int \rho \nabla \Psi \nabla \Psi' + \oint_{\partial \Omega} \rho \alpha \Psi \Psi' \quad (\text{C.1.1})$$

C'est effectivement un produit scalaire si  $\alpha(x) \geq 0$ . En utilisant ce produit scalaire, le gradient de Wasserstein est maintenant

$$-\text{div}(\rho \nabla \Phi).$$

Cependant, comme nous n'avons pas de modèle où la condition (CNH) n'est pas respectée, il demeure quelques questions :

- Avons-nous toujours  $\alpha \geq 0$ ? Si nous n'avons pas  $\alpha \geq 0$  alors (C.1.1) n'est pas un produit scalaire, et notamment le gradient de Wasserstein  $-\text{div}(\rho_0 \nabla \Phi)$  n'est plus une direction de descente, étant donné que l'on peut avoir  $\mathcal{J}(\rho_0^{n+1}) > \mathcal{J}(\rho_0^n)$  pour un pas de descente même infinitésimal.
- Que faire si nous n'avons pas  $\alpha \geq 0$ ?

- Si nous avons  $\alpha \geq 0$ , alors il faut définir une direction de descente. Faut-il prendre les équations des géodésiques classiques pour la descente? Ces équations sont

$$\begin{aligned}\partial_t \rho + \operatorname{div}(\rho v) &= 0 \\ \partial_t v + v \cdot \nabla v &= 0 \\ \rho(0) &= \rho_0, \quad v(0) = -\nabla \Phi.\end{aligned}$$

Mais comme la vitesse n'est plus tangente au bord, les conditions aux bords changent et nous ne savons pas lesquelles prendre à la place.

Les modèles que nous avons utilisés préservent la condition (CNH) (par exemple le modèle de Saint-Venant), mais nous n'avons pas démontré que tous les modèles préservent la condition (CNH). Dans ce cas là, nous avons montré que nous pouvions toujours utiliser le gradient de Wasserstein en utilisant des conditions de Robin (au lieu de Neumann), mais pour lequel il faut que  $\alpha \geq 0$ . Nous n'avons pas réglé le problème du cas possible où  $\alpha \leq 0$ .



# Liste des tableaux

- 2.1 Tableau résumant les différentes écritures de la distance de Wasserstein entre deux densités  $\rho_0$  et  $\rho_1$ , ainsi que quelques propriétés liant les éléments optimaux. . . . . 54
- 6.1 Pour l'expérience d'assimilation de données et les différents algorithmes utilisés, on calcule le temps moyen du calcul de  $\mathcal{W}_2$ , ainsi que l'erreur relative avec l'algorithme de calcul exact, en terme de valeur de  $\mathcal{J}$  et de distance  $\mathcal{L}^2$  à la solution. Le maillage utilisé comporte 100 points. . . . . 135



# Table des figures

1	Contributions des différentes sources d'observations à l'analyse en météorologie. Copyright (2016) IROWG. . . . .	10
1.1	Représentation schématique du 4D-Var . . . . .	21
1.2	Représentation schématique du filtre de Kalman . . . . .	24
1.3	Image de fronts de feu. Source : [Rochoux et al., 2014a]. . . . .	26
1.4	Exemples d'images et de données denses. Copyright (2016) EUMETSAT et NOAA. . . . .	27
1.5	Image de la couverture nuageuse. Copyright (2016) EUMETSAT. . . . .	29
1.6	Distance euclidienne en fonction de la position des carrés . . . . .	30
1.7	Moyenne de deux densités . . . . .	31
1.8	Barycentre de Wasserstein . . . . .	32
2.1	La couverture nuageuse peut être représentée par un élément de $\mathcal{P}(\Omega)$ . (Copyright 2016 Eumetsat). . . . .	35
2.2	Exemple de transport . . . . .	36
2.3	Les plus proches éléments d'une image par rapport à une base de données, pour la distance $L^1$ et la distance de Wasserstein. Source : [Rubner et al., 2000] . . . . .	40
2.4	Étude de la convexité de la distance par rapport au déplacement . . . . .	40
2.5	Exemple de transfert de couleur, où l'on tente de remplacer les couleurs de l'image (a) par celles de l'image (b). Le résultat est l'image (c). Tiré de [Rabin et al., 2011] . . . . .	41
2.6	Moyenne des deux gaussiennes $\rho_0$ et $\rho_1$ pour la distance de Wasserstein et la distance Euclidienne. . . . .	42
2.7	Barycentres de chiffres de formes et de tailles différentes. Source : [Doucet et Cuturi, 2014]. . . . .	42
2.8	Interpolation de texture. Source : [Rabin et Peyré, 2011] . . . . .	43
2.9	Projection d'une densité trouée sur une base de données. Source : [Bonneel et al., 2016] . . . . .	43
2.10	Moyenne de Wasserstein des deux densités $\rho_0$ et $\rho_1$ proches de 1 (voir l'axe des ordonnées). . . . .	44
2.11	Représentation schématique d'un transport $T$ impliqué dans le problème de Monge . . . . .	45
2.12	Représentation schématique d'une mesure de transport pour le problème de Kantorovich . . . . .	49
2.13	Moyenne de Wasserstein $\mathcal{W}_2$ de deux densités, l'une étant en deux parties. . . . .	59
2.14	Différentes moyennes $WF$ de $\rho_0$ et $\rho_1$ pour différentes valeurs de $\delta$ . À comparer avec la figure 2.13 où est tracée la moyenne de Wasserstein. . . . .	60
2.15	Moyenne $WF$ pour des densités de masses différentes. . . . .	60
2.16	Variations des valeurs des distances $\mathcal{L}^2$ , $\mathcal{H}^{-1}$ et $\mathcal{W}_2$ par rapport au déplacement de $\rho_1$ . . . . .	65
3.1	Solution du problème inverse de sismologie en utilisant un coût classique et un coût de Wasserstein. Source : [Métivier et al., 2016]. . . . .	74



3.2	Pour $\rho_0$ et $\rho_1$ on trace le potentiel de Kantorovich $\Phi$ (normalisé) du transport du premier sur le second. Ces données ( $\rho$ et $\Phi$ ) ne sont pas du même type. . . . .	81
3.3	Différence entre la formulation linéaire (3.3.5) qui donne la courbe rouge, et la formulation géodésique (3.3.7) qui donne la courbe verte. . . . .	84
3.4	Courbes de minimisation des algorithmes (DG2), (DG), (DG#) et (DG# $^\nu$ ). . . . .	86
3.5	Valeur d'une fonction coût écrite à l'aide de la distance $\mathcal{H}^{-1}$ en fonction de l'itération et de la méthode utilisée, dépendant du produit scalaire $\mathcal{L}^2$ (12 dans la légende) et $\mathcal{H}^{-1}$ (H-1 dans la légende). Ces méthodes sont la méthode de plus forte pente (gradient dans la légende), gradient conjugué et quasi-Newton. . . . .	90
4.1	Différentes analyses $\rho_0^a$ de la minimisation de $\mathcal{J}(\rho_0) = d_o(\rho_0, \rho_0^b)^2 + \omega_b d_b(\rho_0, \rho_1)^2$ pour plusieurs choix de couples $(d_o, d_b)$ pris parmi les distance $\mathcal{L}^2$ , $\mathcal{H}^{-1}$ et $\mathcal{W}_2$ , et pour plusieurs valeurs de $\omega_b$ . La valeur de $\rho_0^t$ serait quelque part entre l'ébauche et l'observation. . . . .	93
4.2	(a) Courbes d'ébauche, d'état vrai, et d'observations, comparées avec les analyses obtenues en utilisant les différentes distances $\mathcal{L}^2$ et $\mathcal{H}^{-1}$ (b) et $\mathcal{W}_2$ (c), et leur valeur à la sortie du modèle en (d). . . . .	96
4.3	Ébauche, état vrai et valeur de quelques observations utilisées pour contrôler la condition initiale de ce modèle de Shallow-Water. . . . .	97
4.4	Valeurs des analyses de $\mathcal{J}_2$ et $\mathcal{J}_W$ . . . . .	98
4.5	L'état analysé correspondant à la fonction coût $\mathcal{J}_W$ à laquelle on a rajouté une régularisation $\mathcal{L}^2$ à gauche, et $\mathcal{H}^1$ à droite, pour différents poids de régularisation. . . . .	99
4.6	$h(t)$ et $u(t)$ pour un temps $t$ dans la fenêtre d'assimilation, pour l'ébauche, l'état vrai et l'analyse obtenue par une régularisation $\mathcal{H}^1$ pour $\beta = 10^{-2}$ . . . . .	100
4.7	Données utilisées pour l'expérience de la section 4.6. . . . .	100
4.8	Analyse pour la fonction coût $\mathcal{J}_W$ . Celle-ci est placée au milieu, mais sans les détails des deux états utilisés. . . . .	101
4.9	La fonction $\mathcal{J}_2$ arrive bien à retrouver les détails, même si le modèle les diffuse rapidement et sont rapidement invisibles dans les observations, et ce qu'on ait des erreurs de position (à gauche) ou d'amplitude sur l'ébauche (à droite). . . . .	101
4.10	Analyses d'un exemple simple, sans régularisation à gauche (a), avec régularisation à droite (b). . . . .	102
4.11	Observations soumises à du bruit, d'amplitude à gauche (a), et de position à droite (b). . . . .	103
4.12	Analyses pour (a) un coût $\mathcal{L}^2$ et (b) un coût $\mathcal{W}_2$ lorsqu'il n'y a pas de bruit ajouté aux observations, lorsqu'il y a un bruit haute fréquence et lorsqu'il y a un bruit basse fréquence. . . . .	104
4.13	Tracés de l'état vrai, de l'ébauche (avec erreur d'amplitude par rapport à l'état vrai) et des observations (avec erreurs de position par rapport à l'état vrai) utilisées pour l'expérience de la section 4.8. . . . .	105
4.14	Tracés de l'état vrai et de l'ébauche utilisés pour l'expérience d'assimilation à l'aide de la distance $WF$ . L'analyse $h_0^{a,WF}$ est également tracée. . . . .	106
5.1	Variables intervenant dans le modèle de Saint-Venant : au temps $t$ et au point $x$ , $h(t, x)$ est la hauteur de la colonne d'eau et $f(x)$ est la hauteur du fond. La vitesse de l'eau $u(t, x)$ n'est pas représentée ici. . . . .	110
5.2	Éléments utilisés pour la première expérience de contrôle de $u_0$ sans ébauche : $h_0^t$ en haut de (a), $u_0^t$ en bas de (a) et quelques observations en (b). . . . .	113

5.3	Résultats issus des minimisation de $\mathcal{J}_{\mathcal{W}}$ et $\mathcal{J}_2$ , après 30 itérations de chacune des fonctions coût avec le même produit scalaire $\mathcal{L}^2$ . . . . .	114
5.4	Différentes fonctions $h$ et $u$ obtenues pour cette assimilation ayant pour but de contrôler $h_0$ et $u_0$ , dans le cas où il n'y a pas d'ébauche sur $u_0$ . . . . .	115
5.5	Valeurs des différentes fonctions pour cette fois-ci l'assimilation avec un terme d'ébauche sur $u$ . On compare cette figure avec la figure 5.4 pour évaluer l'impact du terme d'ébauche. . . . .	116
5.6	Comparaison de la 9 <sup>ème</sup> observation et de l'état vrai et temps $t_9$ : l'observation contient des erreurs de position. . . . .	116
5.7	Valeurs des différentes fonctions pour cette fois-ci l'assimilation sous erreurs de position sur les observations. Ici encore, seuls $h_0$ et $u_0$ sont contrôlés. . . . .	117
5.8	À gauche (a) : états vrais et quelques observations de la surface libre $h + f$ , à droite (b) : états analysés en fonction des différents algorithmes, et (c) : zoom de la figure (b) en $x = 0.7$ . . . . .	120
5.9	Décroissances de la distance de $f^n$ et $f$ , en fonction de la distance (euclidienne à gauche, Wasserstein à droite), de l'itération $n$ et de l'algorithme utilisé. . . . .	121
5.10	Différents éléments impliqués dans l'expérience de minimisation de $\mathcal{J}_{\mathcal{W}}(h_0, f)$ et $\mathcal{J}_2(h_0, f)$ . En (a) il y a les états vrai, l'ébauche et les deux analyses, en (b) quelques observations, et en (c) la comparaison des analyses et de la dernière observation. . . . .	122
5.11	Analyse trouvée à l'issue de la minimisation de $\mathcal{J}_{\mathcal{W}}$ . . . . .	123
5.12	Éléments utilisés pour l'expérience, à gauche (a) : $h_0$ et $f$ en haut, et $u_0$ en bas ; à droite (b) : trois des 20 observations. . . . .	123
5.13	Analyse trouvée par la minimisation de $\mathcal{J}_{\mathcal{W}}$ , en haut : $h_0$ , au milieu : $u_0$ et en bas : $f$ .124	
6.1	Exemples de traceurs utilisés en océanographie. . . . .	132
6.2	Image d'un tourbillon issue de la plate-forme tournante Coriolis de Grenoble. Source : [Vidard et al., 2014]. . . . .	133
6.3	Ébauche $h_0^b$ , état vrai $h_0^t$ qui a servi à faire les observations via un modèle de Shallow-Water, et les différentes analyses obtenues en fonction des différents algorithmes. Pour ne pas surcharger la légende $h_0^a$ correspond aux analyses de tous les algorithmes utilisés (qui sont proches) sauf $h_0^{a,reg}$ qui correspond à l'analyse issue de la régularisation d'entropie. . . . .	134
6.4	Transport optimal entre la figure de gauche (a) à $t = 0$ et la figure de droite (b) à $t = 1$ , en prenant en compte la côte bretonne, ce qui donne l'interpolation (c) à $t = 1/2$ . Source : [Papadakis et al., 2014]. . . . .	137
6.5	Transport optimal pour l'assimilation de données dans le cas d'occultations. . . . .	139
A.1	À gauche : les croix vertes représentent les points de la grille décalée en temps $\mathcal{G}_d^t$ , et les carrés bleus les points de la grille décalée en espace $\mathcal{G}_d^x$ . À gauche, les points rouges sont sur la grille centrée $\mathcal{G}_c$ . . . . .	144
A.2	Comparaison des algorithmes de descente de gradient simple et de gradient conjugué, pour le calcul du transport optimal entre deux densités $\rho_0$ et $\rho_1$ (en haut). Le gradient conjugué utilise la version "Hesteness-Stiefel" avec réinitialisation de la direction de descente toutes les 100 itérations. En bas à gauche est affichée la valeur de $J$ en fonction des itérations. En bas à droite est tracée la valeur de $\ U^n - U\ $ en fonction de l'itération $n$ , où $U$ est la valeur optimale, obtenue après convergence d'un autre algorithme plus performant (voir A.5). . . . .	146

A.3	Valeur de $\ U^n - U^*\ ^2 + \ V^n - V^*\ ^2$ en fonction de la valeur de $\sigma$ utilisé dans l'algorithme primal dual, pour $n = 500$ . On prend $\tau = 1/\sigma$ . . . . .	150
A.4	Tracé des courbes d'évolution de, <i>en haut à gauche</i> $ L^{n+1} - L^n $ , <i>en haut à droite</i> $ L^n - L^* $ , <i>en bas à gauche</i> $\ G^{n+1} - G^n\ $ et <i>en bas à droite</i> $\ G^n - G^*\ $ , en fonction des itérations, pour trois choix de $\rho_0$ et $\rho_1$ . Les $L^*$ et $G^*$ sont calculés en faisant $10^6$ itérations. . . . .	152
A.5	Comparaison de la vitesse de convergence de $U^n$ vers $U^*$ , c'est-à-dire qu'on trace $\ U^n - U^*\ $ en fonction de l'itération $n$ , pour l'algorithme de descente du gradient conjugué, voir la Section A.3, et pour l'algorithme Primal-Dual. . . . .	153
A.6	Comparaison pour différentes valeurs de $\epsilon$ des barycentres entre les $\rho_0$ et $\rho_1$ . . . . .	154

# Bibliographie

- [Adrian, 1991] ADRIAN, R. J. (1991). Particle-imaging techniques for experimental fluid mechanics. Annual review of fluid mechanics, 23(1):261–304.
- [Agueh, 2008] AGUEH, M. (2008). Gagliardo–Nirenberg inequalities involving the gradient L2-norm. Comptes Rendus Mathématique, 346(13):757–762.
- [Allain, 2016] ALLAIN, A. (2016). Optimal transport without mass conservation : resolution by linear programming and regularization method. Mémoire de D.E.A., Université Grenoble Alpes.
- [Allaire, 2007] ALLAIRE, G. (2007). Numerical analysis and optimization : an introduction to mathematical modelling and numerical simulation. Oxford University Press.
- [Ambrosio et al., 2005] AMBROSIO, L., GIGLI, N. et SAVARÉ, G. (2005). Gradient Flows : In Metric Spaces And In The Space Of Probability Measures. Lectures in math. Birkhauser.
- [Auroux et Fehrenbach, 2011] AUROUX, D. et FEHRENBACH, J. (2011). Identification of velocity fields for geophysical fluids from a sequence of images. Experiments in Fluids, 50(2):313–328.
- [Benamou et Brenier, 1998] BENAMOU, J. et BRENIER, Y. (1998). Weak existence for the semigeostrophic equations formulated as a coupled Monge–Ampère/transport problem. SIAM Journal on Applied Mathematics, 58(5):1450–1461.
- [Benamou, 2003] BENAMOU, J.-D. (2003). Numerical resolution of an “unbalanced” mass transport problem. ESAIM : Mathematical Modelling and Numerical Analysis, 37(05):851–868.
- [Benamou et Brenier, 2000] BENAMOU, J.-D. et BRENIER, Y. (2000). A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische Mathematik, 84(3):375–393.
- [Bigot et Klein, 2015] BIGOT, J. et KLEIN, T. (2015). Characterization of barycenters in the Wasserstein space by averaging optimal transport maps. Preprint.
- [Blayo et al., 2015] BLAYO, E., COSME, E., NODET, M. et VIDARD, A. (2015). Introduction to data assimilation. [http://lgge.osug.fr/meom/pages-perso/cosme/doc\\_cours/ASSIM/Lecture\\_notesAV.pdf](http://lgge.osug.fr/meom/pages-perso/cosme/doc_cours/ASSIM/Lecture_notesAV.pdf).
- [Blayo et al., 2016] BLAYO, E., COSME, E. et VIDARD, A. (2016). Petit manuel d’assimilation de données à l’usage des non-spécialistes. Travail en cours.
- [Bonan, 2013] BONAN, B. (2013). Assimilation de données pour l’initialisation et l’estimation de paramètres d’un modèle d’évolution de calotte polaire. Thèse de doctorat, Université de Grenoble.
- [Bonneel et al., 2016] BONNEEL, N., PEYRÉ, G., CUTURI, M., MAZENC, F., MALISOFF, M., FRIDMAN, E., BONNET, C., MAZENC, F., DJEMA, W., LE HA VY NGUYEN, C. B. et al. (2016). Wasserstein barycentric coordinates : histogram regression using optimal transport. ACM Transactions on Graphics, 35(4).

- [Bormann et al., 2003] BORMANN, N., SAARINEN, S., KELLY, G. et THÉPAUT, J.-N. (2003). The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Monthly Weather Review*, 131(4):706–718.
- [Bouharguane et al., 2014] BOUHARGUANE, A., IOLLO, A. et WEYNANS, L. (2014). Numerical solution of the Monge-Kantorovich problem by Picard iterations. Research Report RR-8477, Inria.
- [Bouharguane et al., 2012] BOUHARGUANE, A., MAITRE, E., OUDET, E. et PAPADAKIS, N. (2012). Multiphysics optimal transportation and image analysis. Preprint.
- [Brenier, 1991] BRENIER, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417.
- [Brenier et al., 2003] BRENIER, Y., FRISCH, U., HÉNON, M., LOEPER, G., MATARRESE, S., MOHAYAEI, R. et SOBOLEVSKIĬ, A. (2003). Reconstruction of the early universe as a convex optimization problem. *Monthly Notices of the Royal Astronomical Society*, 346(2):501–524.
- [Carlier et al., 2015] CARLIER, G., DUVAL, V., PEYRÉ, G. et SCHMITZER, B. (2015). Convergence of entropic schemes for optimal transport and gradient flows. *arXiv preprint arXiv :1512.02783*.
- [Carrillo et al., 2006] CARRILLO, J. A., MCCANN, R. J. et VILLANI, C. (2006). Contractions in the 2-Wasserstein length space and thermalization of granular media. *Archive for Rational Mechanics and Analysis*, 179(2):217–263.
- [Chabot, 2014] CHABOT, V. (2014). *Étude de représentations parcimonieuses des statistiques d’erreur d’observation pour différentes métriques. Application à l’assimilation de données images.* Thèse de doctorat, Université de Grenoble.
- [Chambolle et Pock, 2011] CHAMBOLLE, A. et POCK, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145.
- [Chang et Holt, 1994] CHANG, S. W. et HOLT, T. R. (1994). Impact of assimilating SSM/I rainfall rates on numerical prediction of winter cyclones. *Monthly weather review*, 122(1):151–164.
- [Chizat et al., 2015a] CHIZAT, L., PEYRÉ, G., SCHMITZER, B. et VIALARD, F.-X. (2015a). Unbalanced optimal transport : geometry and Kantorovich formulation. *arXiv preprint arXiv :1508.05216*.
- [Chizat et al., 2016] CHIZAT, L., PEYRÉ, G., SCHMITZER, B. et VIALARD, F.-X. (2016). Scaling algorithms for unbalanced transport problems. *arXiv preprint arXiv :1607.05816*.
- [Chizat et al., 2015b] CHIZAT, L., SCHMITZER, B., PEYRÉ, G. et VIALARD, F.-X. (2015b). An interpolating distance between optimal transport and Fischer-Rao. *arXiv preprint arXiv :1506.06430*.
- [Cordero-Erausquin et al., 2004] CORDERO-ERAUSQUIN, D., NAZARET, B. et VILLANI, C. (2004). A mass-transportation approach to sharp Sobolev and Gagliardo–Nirenberg inequalities. *Advances in Mathematics*, 182(2):307–332.
- [Courtier et al., 1998] COURTIER, P., ANDERSSON, E., HECKLEY, W., VASILJEVIC, D., HAMRUD, M., HOLLINGSWORTH, A., RABIER, F., FISHER, M. et PAILLEUX, J. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I : Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807.
- [Courtier et al., 1994] COURTIER, P., THÉPAUT, J.-N. et HOLLINGSWORTH, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387.
- [Cullen et Gangbo, 2001] CULLEN, M. et GANGBO, W. (2001). A variational approach for the 2-dimensional semi-geostrophic shallow water equations. *Archive for rational mechanics and analysis*, 156(3):241–273.

- [Cuturi, 2013] CUTURI, M. (2013). Sinkhorn distances : Lightspeed computation of optimal transport. In Advances in Neural Information Processing Systems, pages 2292–2300.
- [Deb et al., 2008] DEB, S., KISHTAWAL, C., PAL, P. et JOSHI, P. (2008). A modified tracer selection and tracking procedure to derive winds using water vapor imagers. Journal of Applied Meteorology and Climatology, 47(12):3252–3263.
- [Delon, 2004] DELON, J. (2004). Midway image equalization. Journal of Mathematical Imaging and Vision, 21(2):119–134.
- [Delon et Desolneux, 2010] DELON, J. et DESOLNEUX, A. (2010). Stabilization of Flicker-like effects in image sequences through local contrast correction. SIAM Journal on Imaging Sciences, 3(4): 703–734.
- [Dew et Holmlund, 2000] DEW, G. et HOLMLUND, K. (2000). Investigations of cross-correlation and euclidian distance target matching techniques in the mpef environment. In 5th International Winds Workshop. Lorne, Australia, volume 58.
- [Doucet et Cuturi, 2014] DOUCET, A. et CUTURI, M. (2014). Fast computation of Wasserstein barycenters. In The 31st International Conference on Machine Learning, pages 685–693.
- [Douglas, 2002] DOUGLAS, R. (2002). Rearrangements of functions with applications to meteorology and ideal fluid flow. Large-Scale Atmosphere-Ocean Dynamics I : Analytical methods and numerical models, pages 288–341.
- [Engquist et Froese, 2013] ENGQUIST, B. et FROESE, B. D. (2013). Application of the Wasserstein metric to seismic signals. arXiv preprint arXiv :1311.4581.
- [Engquist et al., 2016] ENGQUIST, B., FROESE, B. D. et YANG, Y. (2016). Optimal transport for seismic full waveform inversion. arXiv preprint arXiv :1602.01540.
- [Esser, 2010] ESSER, J. E. (2010). Primal dual algorithms for convex models and applications to image restoration, registration and nonlocal inpainting. Thèse de doctorat, University of California Los Angeles.
- [Evans, 2010] EVANS, L. (2010). Partial Differential Equations. Graduate studies in mathematics. American Mathematical Society.
- [Farchi et al., 2016] FARCHI, A., BOCQUET, M., ROUSTAN, Y., MATHIEU, A. et QUÉREL, A. (2016). Using the Wasserstein distance to compare fields of pollutants : application to the radionuclide atmospheric dispersion of the Fukushima-Daiichi accident. Tellus B, 68.
- [Ferradans et al., 2014] FERRADANS, S., PAPADAKIS, N., PEYRÉ, G. et AUJOL, J.-F. (2014). Regularized discrete optimal transport. SIAM Journal on Imaging Sciences, 7(3):1853–1882.
- [Feyeux et al., 2016] FEYEUX, N., NODET, M. et VIDARD, A. (2016). Optimal Transport for Data Assimilation. Preprint.
- [Fortin et Glowinski, 1983] FORTIN, M. et GLOWINSKI, R. (1983). Augmented Lagrangian methods : applications to the numerical solution of boundary-value problems. Elsevier.
- [Gramfort et al., 2015] GRAMFORT, A., PEYRÉ, G. et CUTURI, M. (2015). Fast optimal transport averaging of neuroimaging data. In International Conference on Information Processing in Medical Imaging, pages 261–272. Springer.
- [Grossmann et Morlet, 1984] GROSSMANN, A. et MORLET, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM journal on mathematical analysis, 15(4):723–736.
- [Guillemot et Delon, 2016] GUILLEMOT, T. et DELON, J. (2016). Implementation of the midway image equalization. Image Processing On Line, 6:114–129.

- [Hascoët et Pascual, 2004] HASCOËT, L. et PASCUAL, V. (2004). TAPENADE 2.1 user's guide. Technical Report RT-0300, Inria.
- [Henry, 2016] HENRY, M. (2016). Ondelettes et Transport Optimal. Thèse de doctorat, Université Grenoble-Alpes.
- [Hilton et al., 2012] HILTON, F., ARMANTE, R., AUGUST, T., BARNET, C., BOUCHARD, A., CAMY-PEYRET, C., CAPELLE, V., CLARISSE, L., CLERBAUX, C., COHEUR, P.-F. et al. (2012). Hyper-spectral Earth observation from IASI. bulletin of the american meteorological Society, 93(3):347.
- [Hoffman et Grassotti, 1996] HOFFMAN, R. N. et GRASSOTTI, C. (1996). A technique for assimilating SSM/I observations of marine atmospheric storms : tests with ECMWF analyses. Journal of Applied Meteorology, 35(8):1177–1188.
- [Horn et Schunck, 1981] HORN, B. K. et SCHUNCK, B. G. (1981). Determining optical flow. Artificial intelligence, 17(1-3):185–203.
- [Hug, 2016] HUG, R. (2016). Analyse mathématique et convergence d'un algorithme pour le transport optimal dynamique. Thèse de doctorat, Université Grenoble-Alpes.
- [Hug et al., 2015] HUG, R., MAITRE, E. et PAPADAKIS, N. (2015). Multi-physics optimal transportation and image interpolation. ESAIM : Mathematical Modelling and Numerical Analysis, 49(6):1671–1692.
- [Huot et al., 2006] HUOT, E., ISAMBERT, T., HERLIN, I., BERROIR, J.-P. et KOROTAEV, G. (2006). Data assimilation of satellite images within an oceanographic circulation model. In 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, volume 2, pages II–II. IEEE.
- [Ide et al., 1997] IDE, K., COURTIER, P., GHIL, M. et LORENC, A. (1997). Unified notation for data assimilation : operational, sequential and variational. Practice, 75(1B):181–189.
- [Jordan et al., 1998] JORDAN, R., KINDERLEHRER, D. et OTTO, F. (1998). The variational formulation of the Fokker–Planck equation. SIAM journal on mathematical analysis, 29(1):1–17.
- [Kalman, 1960] KALMAN, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of basic Engineering, 82(1):35–45.
- [Kalman et Bucy, 1961] KALMAN, R. E. et BUCY, R. S. (1961). New results in linear filtering and prediction theory. Journal of basic engineering, 83(1):95–108.
- [Kishtawal et al., 2009] KISHTAWAL, C., DEB, S., PAL, P. et JOSHI, P. (2009). Estimation of atmospheric motion vectors from Kalpana-1 imagers. Journal of Applied Meteorology and Climatology, 48(11):2410–2421.
- [Le Dimet et Talagrand, 1986] LE DIMET, F.-X. et TALAGRAND, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. Tellus A, 38(2):97–110.
- [Léonard, 2013] LÉONARD, C. (2013). A survey of the Schrödinger problem and some of its connections with optimal transport. arXiv preprint arXiv :1308.0215.
- [Li et Navon, 2001] LI, Z. et NAVON, I. (2001). Optimality of variational data assimilation and its relationship with the kalman filter and smoother. Quarterly Journal of the Royal Meteorological Society, 127(572):661–683.
- [Lions, 1971] LIONS, J. L. (1971). Optimal control of systems governed by partial differential equations, volume 170. Springer Verlag.

- [Lombardi et Maitre, 2015] LOMBARDI, D. et MAITRE, E. (2015). Eulerian models and algorithms for unbalanced optimal transport. ESAIM : Mathematical Modelling and Numerical Analysis, 49(6):1717–1744.
- [Manobianco et al., 1994] MANOBIANCO, J., KOCH, S., KARYAMPUDI, V. M. et NEGRI, A. J. (1994). The impact of assimilating satellite-derived precipitation rates on numerical simulations of the ERICA IOP 4 cyclone. Monthly weather review, 122(2):341–365.
- [Mariano, 1990] MARIANO, A. J. (1990). Contour analysis : A new approach for melding geophysical fields. Journal of Atmospheric and Oceanic Technology, 7:285–295.
- [Mérigot, 2011] MÉRIGOT, Q. (2011). A multiscale approach to optimal transport. In Computer Graphics Forum, volume 30, pages 1583–1592. Wiley Online Library.
- [Métivier et al., 2016] MÉTIVIER, L., BROSSIER, R., MÉRIGOT, Q., OUDET, E. et VIRIEUX, J. (2016). Measuring the misfit between seismograms using an optimal transport distance : application to full waveform inversion. Geophysical Journal International, 205(1):345–377.
- [Mirouze, 2010] MIROUZE, I. (2010). Régularisation de problèmes inverses à l’aide de l’équation de diffusion, avec application à l’assimilation variationnelle de données océaniques. Thèse de doctorat, Université Paul Sabatier-Toulouse III.
- [Monge, 1781] MONGE, G. (1781). Mémoire sur la théorie des déblais et des remblais. De l’Imprimerie Royale.
- [Nocedal et Wright, 2006] NOCEDAL, J. et WRIGHT, S. J. (2006). Numerical Optimization. Springer series in Operations Research and Financial Engineering. Springer.
- [Otto, 2001] OTTO, F. (2001). The geometry of dissipative evolution equations : the porous medium equation. Communications in Partial Differential Equations, 26:101–174.
- [Otto et Villani, 2000] OTTO, F. et VILLANI, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. Journal of Functional Analysis, 173(2):361–400.
- [Papadakis, 2007] PAPADAKIS, N. (2007). Assimilation de données images : application au suivi de courbes et de champs de vecteurs. Thèse de doctorat, Université Rennes 1.
- [Papadakis et al., 2014] PAPADAKIS, N., PEYRÉ, G. et OUDET, E. (2014). Optimal transport with proximal splitting. SIAM Journal on Imaging Sciences, 7(1):212–238.
- [Percival, 2008] PERCIVAL, J. R. (2008). Displacement Assimilation for Ocean Models. Thèse de doctorat, The University of Reading.
- [Peyré et al., 2012] PEYRÉ, G., FADILI, J. et RABIN, J. (2012). Wasserstein active contours. In 2012 19th IEEE International Conference on Image Processing, pages 2541–2544. IEEE.
- [Piccoli et Rossi, 2014] PICCOLI, B. et ROSSI, F. (2014). Generalized Wasserstein distance and its application to transport equations with source. Archive for Rational Mechanics and Analysis, 211(1):335–358.
- [Piti’e et al., 2005] PITI’E, F., KOKARAM, A. C. et DAHYOT, R. (2005). N-dimensional probability density function transfer and its application to color transfer. In Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1, volume 2, pages 1434–1439. IEEE.
- [Powell, 1964] POWELL, M. J. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. The computer journal, 7(2):155–162.
- [Rabin et Peyré, 2011] RABIN, J. et PEYRÉ, G. (2011). Régularisation de Wasserstein. application au transfert de couleur. In GRETSI’11.



- [Rabin et al., 2011] RABIN, J., PEYRÉ, G., DELON, J. et BERNOT, M. (2011). Wasserstein barycenter and its application to texture mixing. In International Conference on Scale Space and Variational Methods in Computer Vision, pages 435–446. Springer.
- [Ratner et al., 2015] RATNER, V., ZHU, L., KOLESOV, I., NEDERGAARD, M., BENVENISTE, H. et TANNENBAUM, A. (2015). Optimal-mass-transfer-based estimation of glymphatic transport in living brain. In SPIE Medical Imaging, volume 9413, pages 94131J–94131J–6. International Society for Optics and Photonics.
- [Ravela et al., 2007] RAVELA, S., EMANUEL, K. et MCCLAUGHLIN, D. (2007). Data assimilation by field alignment. Physica D : Nonlinear Phenomena, 230(1):127–145.
- [Rochoux et al., 2014a] ROCHOUX, M. et al. (2014a). Vers une meilleure prévision de la propagation d’incendies de forêt : évaluation de modèles et assimilation de données. Thèse de doctorat, Châtenay-Malabry, Ecole centrale de Paris.
- [Rochoux et al., 2014b] ROCHOUX, M. C., RICCI, S., LUCOR, D., CUENOT, B. et TROUVÉ, A. (2014b). Towards predictive data-driven simulations of wildfire spread—part I : Reduced-cost ensemble kalman filter based on a polynomial chaos surrogate model for parameter estimation. Natural Hazards and Earth System Sciences, 14(11):2951–2973.
- [Rubner et al., 2000] RUBNER, Y., TOMASI, C. et GUIBAS, L. J. (2000). The Earth Mover’s distance as a metric for image retrieval. International journal of computer vision, 40(2):99–121.
- [Santambrogio, 2014] SANTAMBROGIO, F. (2014). Introduction to optimal transport theory. Notes.
- [Santambrogio, 2015] SANTAMBROGIO, F. (2015). Optimal transport for applied mathematicians. Progress in Nonlinear Differential Equations and their applications, 87.
- [Sasaki, 1958] SASAKI, Y. (1958). An objective analysis based on the variational method. Journal of the Meteorological Society of Japan. Ser. II, 36(3):77–88.
- [Sasaki et al., 1955] SASAKI, Y., GU, P. et YAN, L. (1955). A fundamental study of the numerical prediction based on the variational principle. J. Meteor. Soc. Japan, 33(6):262–275.
- [Saumier et al., 2010] SAUMIER, L.-P., AGUEH, M. et KHOUIDER, B. (2010). An efficient numerical algorithm for the L2 optimal transport problem with applications to image processing. arXiv preprint arXiv :1009.6039.
- [Schmetz et al., 1993] SCHMETZ, J., HOLMLUND, K., HOFFMAN, J., STRAUSS, B., MASON, B., GAERTNER, V., KOCH, A. et VAN DE BERG, L. (1993). Operational cloud-motion winds from Meteosat infrared images. Journal of applied meteorology, 32(7):1206–1225.
- [Souopgui, 2010] SOUOPGUI, I. (2010). Assimilation d’images pour les fluides géophysiques. Thèse de doctorat, Université Joseph-Fourier-Grenoble I.
- [Tanaka, 1978] TANAKA, H. (1978). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. Probability Theory and Related Fields, 46(1):67–105.
- [Titaud et al., 2010] TITAUD, O., VIDARD, A., SOUOPGUI, I. et LE DIMET, F.-X. (2010). Assimilation of image sequences in numerical models. Tellus A, 62(1):30–47.
- [Uppala et al., 2005] UPPALA, S. M., KÅLLBERG, P., SIMMONS, A., ANDRAE, U., BECHTOLD, V. d., FIORINO, M., GIBSON, J., HASELER, J., HERNANDEZ, A., KELLY, G. et al. (2005). The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society, 131(612):2961–3012.
- [Veersé et Thepaut, 1998] VEERSÉ, F. et THEPAUT, J.-N. (1998). Multiple-truncation incremental approach for four-dimensional variational data assimilation. Quarterly Journal of the Royal Meteorological Society, 124(550):1889–1908.

- [Vidard et al., 2014] VIDARD, A., CHABOT, V. et NODET, M. (2014). Accounting for correlated observation errors in image data assimilation. In Workshop on correlated observation errors in data assimilation, Reading, United Kingdom. ESA - University of Reading.
- [Villani, 2003] VILLANI, C. (2003). Topics in Optimal Transportation. Graduate studies in mathematics. American Mathematical Society.
- [Villani, 2008] VILLANI, C. (2008). Optimal Transport : Old and New. Grundlehren der mathematischen Wissenschaften. Springer.
- [Zhao et al., 2005] ZHAO, Y., WANG, B., JI, Z., LIANG, X., DENG, G. et ZHANG, X. (2005). Improved track forecasting of a typhoon reaching landfall from four-dimensional variational data assimilation of AMSU-A retrieved data. Journal of Geophysical Research : Atmospheres, 110(D14).