# Borne sur l'approximation de noyaux à valeurs opérateurs à l'aide de transformées de Fourier

Romain Brault [1] & Florence d'Alché–Buc [2]

[1] *IBISC, Université d'Évry val d'Essonne,* romain.brault@ibisc.fr.
[2] *LTCI CNRS, Télécom ParisTech,* florence.dalche@telecom-paristech.fr.

**Résumé.** Utilisés pour l'apprentissage multi-tâches et l'apprentissage de sorties structurées, les noyaux à valeurs opérateurs permettent de définir des fonctions à valeurs vectorielles, dans le contexte des espaces de Hilbert à noyaux reproduisants (RKHS). Cependant, les algorithmes d'apprentissage de ces modèles ne passent pas à l'échelle. Dans cet article, il est proposé de définir ces modèles à l'aide de fonctions de re-description aléatoires à valeurs opérateurs en généralisant les fonctions de re-description aléatoires de Fourier proposées par Rahimi et Recht. Le principe de construction de ces fonctions de re-description se fonde sur une généralisation du théorème de Bochner pour les noyaux à valeurs opérateurs de Mercer invariants par translation. Un théorème de convergence uniforme de ce type d'approximation est prouvé pour des noyaux à valeurs opérateurs dont les fonctions de re-description peuvent être non bornées, via l'utilisation d'une inégalité de concentration de Bernstein appropriée aux opérateurs non bornés.

**Mots-clés.** Méthodes à noyaux à valeurs opérateurs, random Fourier features, projections aléatoires, inégalité de concentration.

**Abstract.** Devoted to multi-task learning and structured output learning, operator-valued kernels provide a flexible tool to build vector-valued functions in the context of Reproducing Kernel Hilbert Spaces (RKHS). To scale up these methods, we extend the celebrated Random Fourier Feature methodology to get an approximation of operator-valued kernels. We propose a general principle for Operator-valued Random Fourier Feature construction relying on a generalization of Bochner's theorem for translation-invariant operator-valued Mercer kernels. We prove the uniform convergence of the kernel approximation for bounded and unbounded operator random Fourier features using appropriate Bernstein matrix concentration inequality.

**Keywords.** Operator-valued kernel methods, random Fourier features, random projections, concentration inequality.

# 1 Introduction

L'apprentissage de fonctions à valeurs vectorielles à partir de données est au coeur de la regression multitâche (Evgeniou et collab., 2005), la classification structurée (Dinuzzo

et collab., 2011), l'autoregression à valeurs vectorielles (Lim et collab., 2015) ou encore de l'apprentissage de champs de vecteurs (Baldassarre et collab., 2012). L'intérêt de cette approche réside essentiellement dans la propriété suivante : un modèle à valeurs vectorielles de dimension $p$ peut tirer parti des relations entre les différentes coordonnées du vecteur de sortie afin d'améliorer de manière significative les performances en comparaison à l'apprentissage *indépendant* de $p$ modèles à valeurs scalaires.

Cet article traite des modèles à noyaux à valeurs opérateurs (OVK). Un noyau à valeurs opérateurs sur un espace de Hilbert $\mathcal{Y}$ est défini comme un noyau $\mathcal{Y}$-reproduisant. Tout comme dans le cas scalaire, un noyau à valeurs opérateurs permet de définir un unique espace de Hilbert à noyau reproduisant (RKHS). Pour résoudre un problème d'apprentissage dans cet espace fonctionnel, sous des hypothèses relativement faciles à vérifier, il est possible de s'appuyer sur des théorèmes de représentation pour réduire l'apprentissage dans cet espace fonctionnel à la recherche d'un nombre fini de paramètres, les fonctions solutions s'écrivant uniquement en fonction des valeurs opérateurs de noyaux ancrés sur les données.

Si différents travaux ont montré l'intérêt des modèles à noyaux à valeur opérateurs pour apprendre des fonctions vectorielles, ces derniers souffrent des mêmes maux que leurs cousins scalaires : leur utilisation, dans sa forme naïve, requiert la construction d'une matrice de taille quadratique vis à vis du nombre de données, ainsi que son inversion ou fait appel à des algorithmes de gradient moins coûteux en temps de calcul mais toujours gourmands en mémoire. Face à cette problématique, plusieurs solutions ont étés proposées dans le cas des noyaux à valeurs scalaires. En particulier, les travaux de Rahimi et Recht (2007) ont introduit les re-descriptions aléatoires de Fourier dites *random Fourier features* (RFF) en anglais, largement étudiées, entre autre, par Le et collab. (2013); Sriperumbudur et Szabo (2015); Bach (2015); Sutherland et Schneider (2015). L'approche RFF linéarise un modèle à noyau de type $f(x) = \sum_{i=1}^{n} \alpha_i k(x, x_i)$ en construisant une fonction de re-description explicite mais approchée du type $\tilde{f}(x) = \tilde{\phi}(x)^T \theta$ en exploitant le théorème de Bochner et le fait que tout noyau positif défini invariant par translation peut s'écrire comme la transformée de Fourier d'une mesure. L'approche RFFs s'est montrée compétitive face à d'autres techniques d'approximation comme la méthode de Nÿstrom (Yang et collab., 2012) et a fait ensuite l'objet d'optimisations supplémentaires telles que celle proposée dans *FastFood* de Le et collab. (2013) et de généralisations telles que Li et collab. (2010). Dans ce travail, un principe général de construction d'une fonction de re-description aléatoire de Fourier à valeurs opérateurs est présenté ainsi que l'approximation de noyau correspondante, puis un théorème qui établit la convergence uniforme de l'approximation du noyau vers celui-ci.

# 2 Principe de construction

**Bref rappel des noyaux à valeurs opérateurs**

Soit $\mathcal{Y}$ un espace de Hilbert. Soit $\mathcal{L}(\mathcal{Y})$ l'espace des opérateurs bornés de $\mathcal{Y}$ sur lui-même. Soit $\mathcal{F}(\mathcal{X}, \mathcal{Y})$, l'espace des fonctions de $\mathcal{X}$ dans $\mathcal{Y}$. Dans la suite, $\mathcal{X}$ est un sous-espace de $\mathbb{C}^d$ et $\mathcal{Y}$, un sous-espace de $\mathbb{C}^p$. Dans un but de simplicité, $A^*$ définit l'opérateur adjoint de $A \in \mathcal{L}(\mathcal{Y})$. Un noyau à valeurs opérateurs est définit comme un $\mathcal{Y}$-reproduisant sur $\mathcal{X}$. Une application $K : \mathcal{X} \times \mathcal{X} \to \mathcal{L}(\mathcal{Y})$ est appelé noyau $\mathcal{Y}$-reproduisant si $\sum_{i,j=1}^{N} \langle y_i, K(x_i, x_j) y_j \rangle \geq 0$, pour tout $x_1, \ldots, x_N$ dans $\mathcal{X}$, tout $y_1, \ldots, y_N$ dans $\mathcal{Y}$ et $N \geq 1$.

Soit $x \in \mathcal{X}$, $K_x : \mathcal{Y} \to \mathcal{F}(\mathcal{X}; \mathcal{Y})$ est un opérateur linéaire dont l'action sur un vecteur $y$ est la fonction $K_x y \in \mathcal{F}(\mathcal{X}; \mathcal{Y})$ définie par $(K_x y)(z) = K(z, x)y, \ \forall z \in \mathcal{X}$. De plus, soit un $\mathcal{Y}$-noyau reproduisant $K$, il existe un unique espace de Hilbert $\mathcal{H}_K \subset \mathcal{F}(\mathcal{X}; \mathcal{Y})$ satisfaisant $K_x \in \mathcal{L}(\mathcal{Y}; \mathcal{H}_K), \ \forall x \in \mathcal{X}$ and $f(x) = K_x^* f, \ \forall x \in \mathcal{X}, \forall f \in \mathcal{H}_K$, où $K_x^* : \mathcal{H}_K \to \mathcal{Y}$ est l'opérateur adjoint de $K_x$. L'espace $\mathcal{H}_K$ est alors appelé l'*espace de Hilbert à noyau reproduisant* associé à $K$. Le produit scalaire et la norme induite par cette construction sont notés respectivement $\langle ., . \rangle_K$ et $\|.\|_K$. Les fonctions de $\mathcal{H}_K$ peuvent aussi être définies par une fonction de re-description $\Phi$ appropriée :

**Proposition 2.1 (Carmeli et collab. (2010))** *Soit $\mathcal{H}$ un espace de Hilbert et $\Phi : \mathcal{X} \to \mathcal{B}(\mathcal{Y}; \mathcal{H})$, une fonction de re-description, telle que $\Phi_x \triangleq \Phi(x)$. Alors l'opérateur linéaire $W : \mathcal{H} \to \mathcal{F}(\mathcal{X}; \mathcal{Y})$ définit par $(Wg)(x) = \Phi_x^* g, \ \forall g \in \mathcal{H}, \forall x \in \mathcal{X}$ est une isométrie partielle de $\mathcal{H}$ sur l'espace de Hilbert à noyau reproduisant $\mathcal{H}_K$ ayant pour noyau reproduisant*

$$K(x, z) = \Phi_x^* \Phi_z, \ \ \forall x, z \in \mathcal{X}. \tag{1}$$

**Construction de fonctions de re-description aléatoires à valeurs matricielles**
Un noyau $\mathcal{Y}$-reproduisant est dit $\mathcal{X}$-Mercer lorsque $\mathcal{H}_K$ est un sous espace de $\mathcal{C}(\mathcal{X}; \mathcal{Y})$, espace des fonctions continues de $\mathcal{X}$ dans $\mathcal{Y}$. Il est dit invariant par translation (pour l'addition) si $K(x + a, z + a) = K(x, z)$ pour tout $(x, z, a)$ dans $\mathcal{X}^3$. Un tel noyau est caractérisé par une fonction $K_0 : \mathcal{X} \to \mathcal{L}(\mathcal{Y})$ telle que $K(x, z) = K_0(\delta)$, en prenant $\delta = x - z$. $K_0$ est appelé la *signature* du noyau $K$. Dans le cas des $\mathbb{R}^p$-noyaux de Mercer invariants par translation sur $\mathbb{R}^d$, il est possible de définir un principe de construction d'une fonction de re-description approchée de $K$ en utilisant la propriété suivante prouvée dans Carmeli et collab. (2010) :

**Proposition 2.2 (Instantiation de Carmeli et collab. (2010))** *Soit $\mu$ une mesure sur $\mathbb{R}^d$ et $A : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^p)$ tel que $\langle A(.)y, y' \rangle \in L^1(\mathbb{R}^d, d\mu)$ pour tout couple $y, y' \in \mathbb{R}^p$ et $A(\omega) \geq 0$ pour $\mu$-preque tout $\omega$, alors, la fonction $K_0 : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^p)$ définie pour tout $\delta \in \mathbb{R}^d$, $\forall l, m \in \{1, \ldots, p\}$, par :*

$$[K_0(\delta)]_{lm} = \int_{\mathbb{R}^d} e^{-i\langle \delta, \omega \rangle} \, [A(\omega)]_{lm} \, d\mu(\omega), \tag{2}$$

*i.e. telle que chaque fonction $K_0(\cdot)_{lm}$ est la transformée de Fourier de $A(\cdot)_{lm}$ par rapport à la mesure $\mu(\cdot)$ est la signature d'un noyau $\mathbb{C}^p$-Mercer invariant par translation $K$ sur $\mathbb{R}^d$ tel que $K(x, z) = K_0(x - z)$. De plus, tout noyau $\mathbb{C}^p$-Mercer invariant par translation, est de cette forme pour une paire $(\mathbf{A}(\boldsymbol{\omega}), \boldsymbol{\mu}(\boldsymbol{\omega}))$.*

Remarquons que si en outre, la fonction $A$ et la densité $\nu$ de la mesure (i.e. telle que $d\mu(\omega) = \nu(\omega)d\omega$) sont paires, alors chaque fonction $K_0(\cdot)_{lm}$ est à valeurs réelles et $K_0$ est la signature d'un noyau $\mathbb{R}^p$-Mercer invariant par translation. Dans la pratique plusieurs noyaux de Mercer à valeurs opérateurs ont été utilisés avec succès (voir par exemple, les noyaux décomposables (Micchelli et Pontil, 2005) pour l'apprentissage multi-tâche ou encore les noyaux sans rotationel ou sans divergence appropriés pour l'apprentissage de champs de vecteurs (Baldassarre et collab., 2012)). Il est maintenant possible de définir une fonction de re-description en la choisissant d'après la proposition 2.2.

**Proposition 2.3 (Fonction de re-description de Fourier)** *Soit $B(\omega)$ un opérateur linéaire tel que $A(\omega) = B(\omega)B(\omega)^*$ avec $A(\omega)$ définie comme dans la proposition 2.2. Alors la fonction $\Phi$ définie pour tout $x \in \mathcal{X}$, $\forall (l,m) \in \{1,\dots,p\}^2$, par :*

$$[\Phi(x)]_{lm} = \int_{\mathcal{X}} e^{-i\langle \delta, \omega \rangle} \, [B(\omega)]_{lm} \, d\mu(\omega), \tag{3}$$

*est une fonction de $\mathbb{R}^d$ dans $\mathcal{L}(\mathbb{R}^p)$ de re-description du noyau $\mathbb{R}^p$-Mercer $K$ invariant par translation, i.e. elle satisfait : $\forall x, z \in \mathcal{X}, \Phi(x)^*\Phi(z) = K(x,z)$.*

Sans restriction, la mesure positive $\mu$ peut être vue comme une mesure de probabilité et que sa densité est une densité de probabilité en choisissant une normalisation appropriée : l'équation 3 s'écrit alors : $\Phi(x) = \mathbb{E}_\mu[\exp(-i\langle x, \omega \rangle)B(\omega)^*]$. L'espérance peut être approchée par la méthode de Monte-Carlo et fournir ainsi une approximation $\tilde{\Phi}$ de la fonction de re-description $\Phi$. A l'aide des propriétés trigonométriques, il est possible de définir ainsi une fonction de re-description approchée, $\tilde{\Phi}$, pour tout $x$ dans $\mathcal{X}$, par la concaténation en colonne (notée avec l'opérateur $\oplus$) de toutes les réalisations matricielles $\exp(-i\langle x, \omega_j \rangle)B(\omega_j)^*$ où chaque $\omega_j$ est tiré selon la loi de probabilité $\mu$ :

$$\tilde{\Phi}(x) = \frac{1}{\sqrt{D}} \bigoplus_{j=1}^{D} e^{-i\langle x, \omega_j \rangle} B(\omega_j)^*, \;\; \omega_j \sim \mu, \tag{4}$$

ce qui fournit donc l'approximation $\tilde{S}$ du noyau $K$ : $\forall x, z \in \mathcal{X}$

$$\tilde{S}(x,z) = \tilde{\Phi}(x)^*\tilde{\Phi}(z) = \frac{1}{D} \sum\nolimits_{j=1}^{D} e^{-i\langle x-z, \omega_j \rangle} A(\omega_j).$$

Cette approximation $\tilde{\Phi}(x)^*\tilde{\Phi}(z)$ converge en probabilité vers $\mathbb{E}_\mu \left[\exp(-i\langle x - z, \omega \rangle)A(\omega)\right]$ quand $D$ tend vers l'infini. Le théorème d'inversion proposé par Carmeli et collab. (2010) permet alors de trouver la paire $(\mathbf{A}(\boldsymbol{\omega}), \boldsymbol{\mu}(\boldsymbol{\omega}))$ caractérisant le noyau de Mercer invariant par translation.

**Proposition 2.4 (Carmeli et collab. (2010))** *Soit $K$ un $\mathbb{R}^p$-noyau de Mercer invariant par translation sur $\mathbb{R}^d$ et sa signature $K_0$. Supposons que pour tout $z$ dans $\mathbb{R}^d$ et*

*pour tout $y$, $y'$ dans $\mathbb{R}^p$, $\langle K_0(.)y, y'\rangle \in L^1(\mathbb{R}^d, dx)$ où $dx$ est la mesure de Lebesgue. Posons $C : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^p)$ tel que $\forall \omega \in \mathbb{R}^d$ et pour tout $l, m \in \{1, \ldots, p\}$,*

$$[C(\omega)]_{lm} = \int_{\mathcal{X}} e^{i\langle \delta, \omega\rangle} [K_0(\delta)]_{lm} \, d\delta. \tag{5}$$

*alors*

*i) $C(\omega)$ est un opérateur positif pour tout $\omega \in \mathbb{R}^d$,*

*ii) $\langle C(.)y, y'\rangle \in L^1(\mathbb{R}^d, d\omega)$ pour tout $y$, $y'$ dans $\mathbb{R}^p$,*

*iii) pour tout $\delta \in \mathbb{R}^d$, et pour tout $l, m \in \{1, \ldots, p\}$, $[K_0(\delta)]_{lm} = \int_{\mathcal{X}} e^{-i\langle \delta, \omega\rangle} [C(\omega)]_{lm} \, d\omega$,*

ce qui permet de retrouver $A$ et $\nu$ en remarquant que $C(\omega)d\omega = A(\omega)d\mu(\omega)$.

# 3 Convergence Uniforme

La contribution principale de ce papier est une borne de concentration uniforme de l'approximation $\tilde{S}$ vers le noyau $K$. Celui-ci repose sur l'utilisation d'une inégalité de concentration applicable à des matrices aléatoires non-bornées introduites récemment dans Koltchinskii et collab. (2013). Par la suite, pour une matrice $A$ à coefficients réels, $\|A\|_2$ représente sa norme spectrale tandis que $\|X\|_{\psi_1}$ représente la norme de Orlicz de la variable aléatoire (v.a.) $X$, où $\psi_1 = \exp(|X|) - 1$, permettant de "quantifier la sous-exponentialité" de la v.a. $X$.

**Theorem 3.1 (Convergence uniforme des ORFF)** *Soit $\mathcal{C}$ un compact de $\mathbb{R}^d$ de diamètre $l$. Soit $K$ un $\mathbb{R}^p$-noyau de Mercer invariant par translation sur $\mathbb{R}^d$ et $\nu(\cdot)A(\cdot)$ sa transformée de Fourier inverse dans le sens de la proposition 2.2 avec $\nu(\cdot)$ définissant une densité de probabilité. Soit $D$ un entier positif et $\omega_1, \ldots, \omega_D$, $D$ réalisations indépendantes et identiquement distribuées selon la loi de probabilité $\mu$. Soit $x, z \in \mathcal{C}$. En définissant $F(x - z) = \tilde{S}(x, z) - K(x, z) = \frac{1}{D}\sum_{j=1}^{D} \cos\langle x - z, \omega_j\rangle A(\omega_j) - K(x, z)$, si les constantes suivantes sont finies :*

$$\begin{aligned}
b_D &= \sup_{\delta \in \mathcal{C}} \frac{1}{2}\left( \left\| (K_0(2\delta) + K_0(0)) \, \mathbb{E}_\mu A(\omega) - 2K_0(\delta)^2 \right\|_2 + 2\|\mathbb{V}_\mu A(\omega)\|_2 \right), \\
m &= 4\left( \| \|A(\omega)\|_2 \|_{\psi_1} + \sup_{\delta \in \mathcal{C}} \|K_0(\delta)\| \right), \\
\sigma_p^2 &= \mathbb{E}_\mu \left[ \|\omega\|_2^2 \|A(\omega)\|_2^2 \right],
\end{aligned}$$

*alors pour tout réel $\epsilon$ strictement positif,*

$$\mathbb{P}\left\{ \sup_{\delta \in \mathcal{C}} \|F(\delta)\|_2 \geq \epsilon \right\} \leq C_d \left( \frac{\sigma_p l}{\epsilon} \right)^{\frac{2}{1+2/d}} \begin{cases} \exp\left( -\frac{\epsilon^2 D}{8(d+2)\left(b_D + \frac{\epsilon \bar{u}_D}{6}\right)} \right) & \text{if } \epsilon \bar{u}_D \leq 2(e-1)b_D \\ \exp\left( -\frac{\epsilon D}{(d+2)(e-1)\bar{u}_D} \right) & \text{sinon,} \end{cases}$$

*où,*

$$\bar{u}_D = 2m \log\left(2^{\frac{3}{2}}\left(\frac{m}{b_D}\right)^2\right) \text{ et } C_d = p\left(\left(\frac{d}{2}\right)^{\frac{-d}{d+2}} + \left(\frac{d}{2}\right)^{\frac{2}{d+2}}\right) 2^{\frac{6d+2}{d+2}}.$$

La preuve suit la même structure que la preuve de convergence uniforme dans le cas scalaire présentée dans Rahimi et Recht (2007) et plus récemment détaillée dans Sutherland et Schneider (2015). Dans un premier temps, un $\epsilon$-net est construit sur le compact $\mathcal{C}$. Puis, une inégalité de concentration est appliquée sur chaque point ancre de l'$\epsilon$-net. Enfin, l'erreur commise sur les points hors $\epsilon$-net est majorée grâce à une inégalité de Markov sur la constante de Lipschitz de $F$.

Les approximations des noyaux de Mercer invariants par translation et à valeurs matricielles les plus courants vérifient bien les hypothèses. En particulier, le théorème s'applique aux approximations des noyaux "curl-free" et "divergence-free" issus de noyaux gaussiens utilisés pour l'estimation de champs de vecteurs (voir par exemple Macedo et Castro (2008); Baldassarre et collab. (2012)). Cette approximation ouvre donc la porte à des calculs plus efficaces pour une large gamme de noyaux à valeurs opérateurs.

# Références


Bach, F. 2015, «On the equivalence between quadrature rules and random features», HAl-report-/hal-01118276.

Baldassarre, L., L. Rosasco, A. Barla et A. Verri. 2012, «Multi-output learning via spectral filtering», *Machine Learning*, vol. 87, n° 3, p. 259–301.

Carmeli, C., E. De Vito, A. Toigo et V. Umanità. 2010, «Vector valued reproducing kernel hilbert spaces and universality», *Analysis and Applications*, vol. 8, p. 19–61.

Dinuzzo, F., C. Ong, P. Gehler et G. Pillonetto. 2011, «Learning output kernels with block coordinate descent», dans *Proc. of the 28th Int. Conf. on Machine Learning*.

Evgeniou, T., C. A. Micchelli et M. Pontil. 2005, «Learning multiple tasks with kernel methods», *Journal of Machine Learning Research*, vol. 6, p. 615–637.

Koltchinskii, V. et collab.. 2013, «A remark on low rank matrix recovery and noncommutative bernstein type inequalities», dans *From Probability to Statistics and Back : High-Dimensional Models and Processes–A Festschrift in Honor of Jon A. Wellner*, Institute of Mathematical Statistics, p. 213–226.

Le, Q. V., T. Sarlós et A. J. Smola. 2013, «Fastfood - computing hilbert space expansions in loglinear time», dans *Proceedings of ICML 2013, Atlanta, USA, 16-21 June 2013*, p. 244–252.

Li, F., C. Ionescu et C. Sminchisescu. 2010, *Pattern Recognition : 32nd DAGM Symposium, Darmstadt, Germany, September 22-24, 2010. Proceedings*, chap. Random Fourier Approximations for Skewed Multiplicative Histogram Kernels.

Lim, N., F. d'Alché-Buc, C. Auliac et G. Michailidis. 2015, «Operator-valued kernel-based vector autoregressive models for network inference», *Machine Learning*, vol. 99, n° 3, p. 489–513.

Macedo, Y. et R. Castro. 2008, «Learning div-free and curl-free vector fields by matrix-valued kernels», cahier de recherche, Preprint A 679/2010 IMPA.

Micchelli, C. A. et M. A. Pontil. 2005, «On learning vector-valued functions», *Neural Computation*, vol. 17, p. 177–204.

Rahimi, A. et B. Recht. 2007, «Random features for large-scale kernel machines», dans *NIPS 20, Vancouver, British Columbia, Canada, December 3-6, 2007*, p. 1177–1184.

Sriperumbudur, B. et Z. Szabo. 2015, «Optimal rates for random fourier features», dans *Advances in Neural Information Processing Systems 28*, édité par C. Cortes, N. Lawrence, D. Lee, M. Sugiyama et R. Garnett, p. 1144–1152.

Sutherland, D. J. et J. G. Schneider. 2015, «On the error of random fourier features», dans *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, UAI 2015, July 12-16, 2015, Amsterdam, The Netherlands*, p. 862–871.

Yang, T., Y.-F. Li, M. Mahdavi, R. Jin et Z. Zhou. 2012, «Nyström method vs random fourier features : A theoretical and empirical comparison», dans *NIPS 25*, édité par F. Pereira, C. Burges, L. Bottou et K. Weinberger, p. 476–484.