Solving linear equations with messenger-field and conjugate gradient techniques: An application to CMB data analysis

Jan Papez, Laura Grigori, Radek Stompor

Solving linear equations with messenger-field and conjugate gradients techniques – an application to CMB data analysis

Jan Papež, Laura Grigori, Radoslav Stompor

Solving linear equations with messenger-field and conjugate gradients techniques – an application to CMB data analysis

Jan Papež[*][†], Laura Grigori[*], Radoslav Stompor[‡]

Project-Team Alpines

Abstract: We discuss linear system solvers invoking a messenger-field and compare them with (preconditioned) conjugate gradients approaches. We show that the messenger-field techniques correspond to fixed point iterations of an appropriately preconditioned initial system of linear equations. We then argue that a conjugate gradient solver applied to the same preconditioned system, or equivalently a preconditioned conjugate gradient solver using the same preconditioner and applied to the original system, will in general ensure at least a comparable and typically better performance in terms of the number of iterations to convergence and time-to-solution. We illustrate our conclusions on two common examples drawn from the Cosmic Microwave Background data analysis: Wiener filtering and map-making. In addition, and contrary to the standard lore in the CMB field, we show that the performance of the preconditioned conjugate gradient solver can depend importantly on the starting vector. This observation seems of particular importance in the cases of map-making of high signal-to-noise sky maps and therefore should be of relevance for the next generation of CMB experiments.

Key-words: Numerical methods - linear systems solvers; cosmic microwave background data analysis - Wiener Filter - map-making

---

[*] INRIA Paris, Sorbonne Université, Univ Paris-Diderot SPC, CNRS, Laboratoire Jacques-Louis Lions, équipe ALPINES, France
[†] corresponding author: `Jan.Papez@inria.fr`
[‡] AstroParticule et Cosmologie, Univ Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs de Paris, Sorbonne Paris Cité, France

Résolution d'équations linéaires avec un champ messager et des techniques de gradients conjugués - une application à l'analyse de données en CMB

Résumé : Nous discutons la résolution de systèmes linéaires utilisant un champ messager et comparons cette approche avec la méthode du gradient conjugué (préconditionné). Nous montrons que les techniques du champ messager correspondent à des itérations de point fixe d'un système d'équations linéaires préconditionné de manière approprié. Nous argumentons ensuite qu'un solveur de gradient conjugué appliqué au même système préconditionné, ou de manière équivalente un solveur de gradient conjugué préconditionné utilisant le même préconditionneur et appliqué au système original, assurera en général au moins une performance comparable et typiquement meilleure en termes de nombre d'itérations pour converger et de délai de résolution.

Mots-clés : Méthodes numériques, résolution de systèmes linéaires, analyse du fond difus cosmique

1   Introduction.

Studies of the Cosmic Microwave Background (CMB) anisotropies have been driving the progress in our understanding of the Universe for nearly quarter of the century. The current forefront of the CMB research is the characterization of polarization properties of the CMB anisotropies. The next generation of the CMB observatories has been, and is, designed to ensure that the scientific potential of this new probe is fully exploited. This will call for advanced, high performance data analysis techniques applicable to enormous data sets which will be collected by these new observatories.

CMB data analysis commonly involves solutions of large, structured linear systems of equations. Two typical and important examples of such systems are map-making and Wiener-Filter systems of equations (see, e.g., [10, 1], respectively, for early pioneering work and [14, 18] for examples of more recent applications). These systems are solved either as a stand-alone task or as part of more involved process, such as a power spectrum estimation, which commonly require multiple solutions of such systems. In this work we study, from theoretical and practical perspective, two specific algorithms for solving such systems of equations: a preconditioned conjugate gradient (PCG) approach and a messenger-field (MF) technique. Both these approaches have been applied in the context of the applications considered here. Of the two, the PCG approach has been more popular and more broadly used to date. Nevertheless, it has been argued in a number of recent papers (e.g., [4, 16, 9]) that the messenger-field approach can be highly efficient for both these applications and can deliver performance in some cases exceeding that of some specific PCG approaches, while at the same time be more generally feasible, and straightforward to implement and apply [16, 9, 8]. We note that this is the combination of all these features, which makes the MF approach potentially attractive. Indeed, performance of the PCG technique is hinged on a choice of a preconditioner matrix, $\mathbf{M}$, and while very efficient preconditioners can be constructed, in principle outperforming other methods, the construction typically gets quickly difficult and potentially prohibitive from the computational point of view.

To be specific let us consider a linear system of equations,

$$\mathbf{A}\,\mathbf{x} \;=\; \mathbf{b} \tag{1}$$

where the system matrix $\mathbf{A}$ is symmetric and positive definite (SPD). Instead of solving this equation directly, in the PCG approach one solves the preconditioned system,

$$\mathbf{M}^{-1}\,\mathbf{A}\,\mathbf{x} \;=\; \mathbf{M}^{-1}\,\mathbf{b}, \tag{2}$$

applying the conjugate gradient (CG) technique [5]. If the preconditioner is chosen in the way that $\mathbf{M}^{-1}\mathbf{A}$ is better conditioned than the original system matrix $\mathbf{A}$, then the solution can be derived in a fewer and often significantly fewer iterations. Hereafter, we define the condition number κ as

$$\kappa(\mathbf{A}) \;\equiv\; \|\mathbf{A}\|_2 \cdot \|\mathbf{A}^{-1}\|_2 \tag{3}$$

where $\|\mathbf{A}\|_2$ is the spectral norm[1] of the matrix $\mathbf{A}$. For a good preconditioner $\kappa(\mathbf{A}) \gg \kappa(\mathbf{M}^{-1}\mathbf{A}) \geq 1$. A preconditioner is then better, when its inverse succeeds in capturing more essential features of the inverse system matrix $\mathbf{A}^{-1}$, which can not be computed directly by assumption. Consequently, a better preconditioner will typically be computationally more demanding, requiring potentially some precomputation in order to be constructed, as well as more floating point operations on each iteration step whenever its inverse has to be applied to a vector (see e.g., [6, 21] for a relevant example from the CMB map-making). Thus the number of iterations needed to achieve some desired precision of the solution may not always be a sufficient performance comparison metric.

While these observations make the method comparison cumbersome and potentially limited to very specific cases and concrete implementations, the question which class of methods is more promising in ensuring performance sufficient for forthcoming data sets, e.g., in the context of the CMB field, is of actual practical importance. This is the question we tackle in this work.

Our methodology is as follows. We first show that any MF method applied to a linear system involves preconditioning of the original set of equations with a specific preconditioner. Then we argue on theoretical grounds, and

---

[1]The spectral norm is equal to the largest singular value of the matrix. If the matrix is symmetric, the norm is equal to the largest (in magnitude) eigenvalue and the condition number to the ratio of the largest and the smallest eigenvalues.

later demonstrate on a few study cases that the corresponding PCG algorithm is at least as efficient as, and often much better than, the MF technique, while featuring similar computational complexity. Combining the MF method with a so-called cooling technique can further improve its performance at least in some cases. Nonetheless, this does not seem to affect the overall assessment at least for the specific cooling prescriptions studied in this paper and motivated by earlier work. Consequently, with an exception of some particular cases, where the MF technique may be found more suitable, in general the PCG approach is more promising and should be the method of choice. Though we demonstrate our conclusions using specific examples from applications to the Wiener filter and map-making procedures, we expect them to hold more generally.

This paper is organized as follows. In Sect. 2 we provide a general discussion of the messenger-field technique and contrast it with the PCG solvers. In Sec. 3 we discuss some general numerical aspects of their implementations and performance, which we later illustrate with help of numerical experiments applied to simulated CMB data and involving applications of both these techniques to polarized Wiener filtering, Sect. 4, and map-making, Sect. 5. We conclude in Sect. 6. Some more technical considerations are deferred to the Appendices.

2 Messenger-field iterative solver.

In this Section we first introduce the messenger-field solvers in a consistent and general algebraic framework and outline its relation with the PCG method. We specialize this general discussion in the follow-up Sections.

2.1 Basic approach.

Let us consider a system of linear equations as in Eq. (1). In general, the messenger-field approach involves a split of the system matrix $\mathbf{A}$, such as

$$\mathbf{A} \equiv \mathbf{C} - \mathbf{D} = \mathbf{C}\left(\mathbf{I} - \mathbf{C}^{-1}\mathbf{D}\right), \tag{4}$$

where $\mathbf{C}$ is invertible by construction and its inverse is easy to compute. $\mathbf{I}$ is an identity matrix. After multiplying Eq. (1) by $\mathbf{C}^{-1}$ from the left (which corresponds to *preconditioning* the original system), we get the system

$$\left(\mathbf{I} - \mathbf{C}^{-1}\mathbf{D}\right)\mathbf{x} = \mathbf{C}^{-1}\mathbf{b}. \tag{5}$$

The MF method introduces an extra data object $\mathbf{t}$, the messenger field, which can be defined as

$$\mathbf{t} \equiv \mathbf{D}\mathbf{x} + \mathbf{b}, \tag{6}$$

so Eq. (5) can be represented as

$$\mathbf{t} = \mathbf{D}\mathbf{x} + \mathbf{b}, \tag{7}$$
$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{t}. \tag{8}$$

This can be used to define an iterative scheme

$$\mathbf{t}^{(i+1)} = \mathbf{D}\mathbf{x}^{(i)} + \mathbf{b}, \tag{9}$$
$$\mathbf{x}^{(i+1)} = \mathbf{C}^{-1}\mathbf{t}^{(i+1)}. \tag{10}$$

We note that the messenger field $\mathbf{t}$ introduced in this way is a dummy object. Therefore, barring some implementational advantages, the equations above are equivalent to a reduced system from which the messenger field has been explicitly eliminated and which can be directly derived from Eq. (5). This reads,

$$\mathbf{x} = \mathbf{C}^{-1}\mathbf{D}\mathbf{x} + \mathbf{C}^{-1}\mathbf{b}, \tag{11}$$

and the corresponding iterative scheme is given by

$$\mathbf{x}^{(i+1)} = \mathbf{C}^{-1}\mathbf{D}\mathbf{x}^{(i)} + \mathbf{C}^{-1}\mathbf{b}. \tag{12}$$

This is a *fixed-point iteration* scheme (e.g., [17]) and its derivation is analogous to the derivation of the classical iteration methods that also rely on the splitting of the system matrix as in Eq.(4). The Jacobi iterative method takes $\mathbf{C}$ as the diagonal of $\mathbf{A}$, while in the Gauss–Seidel method $\mathbf{C}$ is equal to the lower triangular part (including the diagonal) of $\mathbf{A}$. Therefore, the convergence results of the classical iteration methods (see, e.g., [17, Sections 4.1–4.2], or [5, Section 10.1]) can be applied also for the messenger field. In particular, the iterations in Eq. (12) will always (asymptotically) converge to the true answer provided that the absolute value of the eigenvalues of $\mathbf{C}^{-1}\mathbf{D}$ is bound to be smaller than unity. Indeed, it can be seen from Eqs. (11) and (12) that the error of the ith approximation satisfies

$$
\begin{aligned}
\epsilon^{(i)} &\equiv \mathbf{x} - \mathbf{x}^{(i)} = \mathbf{C}^{-1}\mathbf{D}\left(\mathbf{x} - \mathbf{x}^{(i-1)}\right) \\
&\equiv \mathbf{C}^{-1}\mathbf{D}\,\epsilon^{(i-1)} = \left[\mathbf{C}^{-1}\mathbf{D}\right]^{i}\epsilon^{(0)}
\end{aligned}
\tag{13}
$$

This is the case of the splits used in the Wiener Filter and map-making examples considered in this paper as shown in Appendix B. The convergence of the method will depend on the specific values of the eigenvalues and the eigenmode with the largest eigenvalue, i.e., closest to 1, will be the slowest to converge.

We emphasize that whether we choose to implement the single equation version as in Eq. (12) or the double equation one as in Eqs. (9)–(10), the result will be the same to within numerical precision as both these cases we solve the same linear system, Eq. (5), performing equivalent iterations. Consequently, the messenger-field approach is a fixed point iteration technique applied to a preconditioned system in Eq. (5). However, this equation can be solved using other means, for instance, a conjugate gradient (CG) approach, which is typically more efficient than the fixed point iterations (see, e.g., [11, Section 5.5] as well as Sect. 3.1). Moreover, solving Eq. (5) with help of the CG technique is equivalent to solving the initial set of equations, Eq. (1), using a PCG technique with the preconditioner set to $\mathbf{M} \equiv \mathbf{C}$. In cases when the fixed point method is expected to converge very efficiently, i.e., when $\mathbf{A} \simeq \mathbf{C}$, also the PCG solver will perform well since $\mathbf{C}^{-1}\mathbf{A} \simeq \mathbf{I}$, a hallmark of a good preconditioner. Similarly, the MF solver based on the split involving a good preconditioner will likely be efficient. From the computational point of view, both techniques require multiple applications of the inverse preconditioner $\mathbf{M}^{-1}$ to a vector, thus resulting in similar numerical cost.

The main message of this Section is that the messenger-field method involves fixed point iterations applied to a preconditioned system of linear equations. Its performance is determined by an adopted split of the system matrix, which also defines the preconditioner applied to precondition the initial system. This preconditioner can be used alternately in a PCG solver employed to solve directly the initial system, what is expected to ensure performance as good as, or better than, that of the MF technique as far as the number of iterations but also time to convergence are concerned.

2.2 Cooling technique.

The convergence of the fixed-point method, Eq. (13), depends on the components of the initial error $\mathbf{x} - \mathbf{x}^{(0)}$ in the invariant subspaces associated with the eigenvalues of $\mathbf{C}^{-1}\mathbf{D}$, especially with the dominant (largest) ones. The cooling technique proposed in [4] aims at providing iteratively a good initial guess $\mathbf{x}^{(0)}$. In the general setting considered above, the cooling technique replaces the original problem, Eq. (1), represented in a split form as in Eq. (4) by

$$
\mathbf{A}(\lambda)\mathbf{x}(\lambda) = \mathbf{b}(\lambda), \qquad \mathbf{A}(\lambda) = \mathbf{C}(\lambda) - \mathbf{D}(\lambda)
\tag{14}
$$

where the *cooling parameter* λ is defined so that: (a) for $\lambda = 1$, the above problem is equivalent to the original problem in Eqs. (1) and (4); (b) for $\lambda \to \infty$, $\mathbf{D}(\lambda) \to 0$ and $(\mathbf{C}(\lambda))^{-1}\mathbf{D}(\lambda) \to 0$. This means that for large values of λ an accurate solution of the modified system, Eq. (14), i.e, $\mathbf{x}(\lambda)$, can be recovered within few (fixed-point) iterations. Naturally, $\mathbf{x}(\lambda)$ can be far from the desired solution $\mathbf{x}(1)$, however, it is claimed that $\mathbf{x}(\lambda)$ is a good starting vector for fixed-point iterations corresponding to the modified problem with a smaller value λ. In the cooling method λ is progressively adapted in the course of the iterations with its value gradually decreasing from an initial, and rather large value, down to 1. While no general prescription is given in the literature, it has been claimed [4, 9] that at least in some applications significant gains can be derived as compared to the fixed point iterations, if the rate of change of λ is appropriately tuned. In most of the studied applications the value of λ is kept constant for some number of iterations and after that it is discretely changed to a lower value. The relative comparison of the performance of the cooling method with that

of the PCG solver of the initial equation, i.e., with $\lambda = 1$, is unclear and the freedom in defining the rate at which λ is changed makes the mathematical analysis of this method and in particular its potential advantages over others difficult. Consequently, in this work we resort to numerical experiments to investigate pros and cons of this technique in the specific cases of interest, Sects. 4 and 5.

We note however that a PCG solver could be used instead of the fixed point iterations within the cooling scheme. Though, the fixed point iterations would still be preferable whenever the value of λ is adjusted after each or every few iterations, e.g., as in the cooling scheme proposed in [4, Sect. 2.2]. However, in the cases when the value of λ is kept unchanged over a number of iterations, as in the numerical experiments presented in [9] and in [16], replacing the fixed-point iterations by a PCG method is expected to result in some performance gain accumulated from all the gains obtained from the solves for a fixed value of λ.

3 Numerical considerations.

3.1 Convergence.

Given a current approximation, $\mathbf{x}^{(i)}$, to the solution of Eq. (1), the appropriate measure of the error $\mathbf{x} - \mathbf{x}^{(i)}$, which has in most of the applications a physical meaning, is given by the $\mathbf{A}$-norm (often called *energy norm)* of the error,

$$\|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}} \equiv \left((\mathbf{x} - \mathbf{x}^{(i)})^T \mathbf{A} (\mathbf{x} - \mathbf{x}^{(i)}) \right)^{1/2}. \tag{15}$$

We note that the energy norm of the error requires the exact solution of Eq. (1) and it is therefore not computable. In practice, it is often replaced by other choices such as the relative (euclidean) norm of the residual, which however may give misleading information about the convergence in cases when matrix $\mathbf{A}$ is ill-conditioned. We discuss some alternative options later on.

We now show that the PCG method is superior to the fixed-point method in terms of minimizing the energy norm of the error. We first recall key properties of (P)CG approach; see, e.g., [17, Lemma 6.28].

Let $\mathbf{x}^{(CG,i)}$ be the ith approximation given by the CG method for solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ with the initial guess $\mathbf{x}^{(0)}$. Then,

$$\mathbf{x} - \mathbf{x}^{(CG,i)} = \hat{\varphi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)}), \tag{16}$$

where $\hat{\varphi}_i$ is a polynomial with $\deg(\hat{\varphi}_i) \le i$, $\hat{\varphi}_i(0) = 1$, what we write succinctly as $\hat{\varphi}_i \in \mathcal{P}_0^i$, and,

$$\|\hat{\varphi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \min_{\hat{\psi}_i \in \mathcal{P}_0^i} \|\hat{\psi}_i(\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}}. \tag{17}$$

Similarly, when $\mathbf{x}^{(PCG,i)}$ is the ith approximation given by the PCG method for solving the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ preconditioned by $\mathbf{C}$, using the initial guess $\mathbf{x}^{(0)}$, we have

$$\mathbf{x} - \mathbf{x}^{(PCG,i)} = \varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)}), \tag{18}$$

with $\varphi_i \in \mathcal{P}_0^i$ and

$$\|\varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \min_{\psi_i \in \mathcal{P}_0^i} \|\psi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}}. \tag{19}$$

Let us now consider the fixed-point method as defined in Eq. (12) assuming the same initial guess, $\mathbf{x}^{(0)}$. From Eqs. (19) and (13),

$$\|\mathbf{x} - \mathbf{x}^{(PCG,i)}\|_{\mathbf{A}} = \|\varphi_i(\mathbf{C}^{-1}\mathbf{A})(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} \tag{20}$$

$$\le \|(\mathbf{C}^{-1}\mathbf{D})^i(\mathbf{x} - \mathbf{x}^{(0)})\|_{\mathbf{A}} = \|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}}, \tag{21}$$

as $(\mathbf{C}^{-1}\mathbf{D})^i = (\mathbf{I} - \mathbf{C}^{-1}\mathbf{A})^i$ and $\psi_i(x) \equiv (1 - x)^i \in \mathcal{P}_0^i$. This means that, in terms of the energy norm of the error, the PCG method converges at least as fast as the fixed-point method. In practice, one can however expect significantly faster convergence, as suggested by Eq. (19).

Finally, we recall that the (preconditioned) conjugate gradient method yields a *monotonic* decrease of the energy norm of the error, i.e.,

$$\|\mathbf{x} - \mathbf{x}^{(PCG,i+1)}\|_{\mathbf{A}} \leq \|\mathbf{x} - \mathbf{x}^{(PCG,i)}\|_{\mathbf{A}}, \qquad i = 0, 1, \ldots. \tag{22}$$

This does not hold, in general, for the approximations given by the fixed-point iterations. Even for splitting of $\mathbf{A}$ such that the largest eigenvalue of $\mathbf{C}^{-1}\mathbf{D}$ is strictly smaller than 1, and therefore $\mathbf{x}^{(i)} \to \mathbf{x}$, $i \to \infty$, one can observe (especially in initial iterations) that

$$\|\mathbf{x} - \mathbf{x}^{(i+1)}\|_{\mathbf{A}} > \|\mathbf{x} - \mathbf{x}^{(i)}\|_{\mathbf{A}}. \tag{23}$$

3.2 Computational complexity.

Regarding the computational and memory cost of each iteration, the fixed point scheme, Eq. (12), is the simplest and the cheapest method as it requires to evaluate $\mathbf{C}^{-1}\mathbf{D}\mathbf{x}^{(i)}$ only once per iteration and to store only two vectors, $\mathbf{x}^{(i)}$, $\mathbf{C}^{-1}\mathbf{b}$. The PCG method requires more memory needed to store up to 4 or 5 vectors, depending on the implementation and each iteration requires additionally two inner products plus some scalar multiplications and vector updates. However, typically and in particular in the applications considered in this paper, the most time consuming operations are the multiplications by matrix $\mathbf{A}$ and by $\mathbf{C}^{-1}$, rendering these additional costs mostly irrelevant. As an example in Appendix A we describe an implementation of the PCG algorithm in the context of the Wiener Filter that allows to perform a single PCG iteration with the computational cost comparable to the cost of one fixed-point iteration, Eq. (12).

We can further capitalize on using the PCG method whenever the relative residual or an error measure corresponding to the $\mathbf{A}$-norm of the error need to be frequently evaluated; in the extreme case, at each iteration. The residual $\mathbf{r}^{(i)}$ is updated on each PCG iteration and it is therefore at our disposal; this is not the case of the fixed-point iterations, Eq. (12). Similarly, there is a numerically stable way to evaluate the problem-related error measure corresponding to the $\mathbf{A}$-norm of the error, see also Appendix A. This evaluation involves only scalar quantities that are already at our disposal during the PCG iterations.

We conclude that in terms of the time per iterations, both approaches, the MF and the corresponding PCG, are comparable, and therefore the number of iterations to convergence is a sufficient comparison metric.

4 Application to Wiener Filtering.

4.1 The problem.

Let us consider a sky map $\mathbf{m}$ composed of a sky signal $\mathbf{s}$ and some noise $\mathbf{n}$ due to our instrument, so

$$\mathbf{m} = \mathbf{s} + \mathbf{n}. \tag{24}$$

We assume that the sky signal is Gaussian over an ensemble of sky realizations with zero mean and known covariance given by $\mathbf{S}$. The noise is also Gaussian with zero mean and the covariance given by $\mathbf{N}$ over the ensemble of noise realizations. We further assume that the noise is uncorrelated and therefore its covariance $\mathbf{N}$ is block-diagonal. The minimum variance estimate of the sky signal, i.e., its Wiener Filter, is then given by e.g., [1],

$$\mathbf{s}_{WF} = \left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)^{-1} \mathbf{N}^{-1} \mathbf{m}. \tag{25}$$

Computing the Wiener Filter of the measured map, $\mathbf{m}$, does require an inversion of the system matrix, $\mathbf{S}^{-1} + \mathbf{N}^{-1}$. As modern CMB maps may contain up to many millions of pixels this task can indeed be daunting. This is because in general there is no obvious domain in which both the signal and noise covariances are simultaneously diagonal. Indeed, the signal covariance $\mathbf{S}$ is diagonal in harmonic domain, where the pixel-domain map $\mathbf{m}$ is described by a vector of coefficients $\mathbf{m}_{\ell m}$ obtained as a result of a spherical harmonic transform applied to the map, while the noise covariance is diagonal in the pixel domain and only diagonal in the harmonic one, if the noise is homogeneous what

is unlikely to be realized in practice. Consequently, a standard way to tackle this problem is to rewrite Eq. (25) as a linear set of equations,

$$\left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)\mathbf{s}_{WF} = \mathbf{N}^{-1}\mathbf{m}, \tag{26}$$

and solve it using some iterative method (e.g., [19]). Both CG and PCG techniques have been applied in this context and while the former was found to show a rather unsatisfactory convergence rate, it was demonstrated that it could be improved significantly albeit with help of a rather advanced and involved from the implementation point of view preconditioner borrowed from multi-grid techniques [19].

The MF method, originally proposed in this context by [4], involves splitting the noise covariance into homogeneous and inhomogeneous parts by representing $\mathbf{N} \equiv \bar{\mathbf{N}} + \mathbf{T}$, where $\mathbf{T} = \tau\mathbf{I}$, is a homogeneous part, $\tau = \min(\mathrm{diag}(\mathbf{N}))$. This leads to a split of the system matrix $\mathbf{S}^{-1} + \mathbf{N}^{-1}$ owing to the fact that,

$$\mathbf{N}^{-1} = \mathbf{T}^{-1} - \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}; \tag{27}$$

see, e.g., [7, p. 258]. Taking then,

$$\mathbf{C} \equiv \mathbf{S}^{-1} + \mathbf{T}^{-1}, \tag{28}$$

$$\mathbf{D} \equiv \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}, \tag{29}$$

and on introducing the messenger field $\mathbf{t}$, Eq. (6), we can rewrite Eq. (25) in its messenger-field representation, i.e.,

$$
\begin{aligned}
\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)\mathbf{t}^{(i+1)} &= \mathbf{T}^{-1}\mathbf{s}_{WF}^{(i)} + \bar{\mathbf{N}}^{-1}\mathbf{m}, \\
\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)\mathbf{s}_{WF}^{(i+1)} &= \mathbf{T}^{-1}\mathbf{t}^{(i+1)},
\end{aligned}
\tag{30}
$$

with the former equation solved in the pixel and the latter in the harmonic domain and with the spherical harmonic transforms used to switch between these domains. These equations are equivalent to Eqs. (3) and (4) of [4]. Their numerical experiments showed that the solver tended to converge quickly to the solution given desired precision and therefore the method was proposed as an efficient way to resolve the slow convergence problem of the CG method without a need for potentially complex preconditioners needed for an efficient PCG solver where both these methods should be applied directly to the initial problem, Eq. (26).

As argued earlier, Eqs. (30) are equivalent to a fixed-point iteration solver applied to the equation, Eq. (5),

$$
\left(\mathbf{I} - \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}\right)\mathbf{s}_{WF} =
$$
$$
= \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\bar{\mathbf{N}}^{-1}\mathbf{m}, \tag{31}
$$

what can be rewritten in an explicitly iterative form as,

$$\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)\mathbf{s}_{WF}^{(i+1)} = \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}\mathbf{s}_{WF}^{(i)} + \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\bar{\mathbf{N}}^{-1}\mathbf{m}. \tag{32}$$

In the next Section, we compare the performance of different solvers applied to Eqs. (26), (31), and (32). From the general consideration of the previous section our expectation is that the CG solver applied to Eq. (31), and equivalent to the PCG solution of Eq. (26) with $\mathbf{M} \equiv \mathbf{C} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$ should perform better than the messenger-field solver, Eq. (32).

4.2 Simulated cases.

To demonstrate and validate our analytical expectation we apply both these solvers to simulated data sets. These are obtained as follows. We first generate maps of three Stokes parameters, I, Q and U, in the Healpix pixelization with the Healpix resolution parameter n_{side} set to 512. These maps are computed using a HEALpy routine, `synfast`, providing as the input CMB power spectra as computed for the standard cosmological model with the current best values of

the parameters [13]. In the following calculations we set the band-limit of the sky signal ℓ_{max} to $2n_{side}$. This is low enough to ensure the orthogonality of the relevant spherical harmonics over the grid of Healpix pixels. However, it leads to a rank-deficient signal covariance matrix. Consequently, hereafter, its inverse, $\mathbf{S}^{-1}$, is to be understood as a pseudo-inverse. We checked on selected cases that setting ℓ_{max} to $3n_{side}$, did not impact our conclusions. We add to these sky maps inhomogeneous albeit uncorrelated noise with rms changing over the sky as in the case of the WMAP observations[2].

We consider two cases with either full or partial sky observations. In this latter case, only 20% of the sky is observed corresponding to the polar cap regions as defined by the Planck HFI mask[3].

4.3 Numerical results.

We consider hereafter the following solvers:

- CG applied to the redefined system, Eq. (31), that is equivalent to PCG applied to the original system Eq. (26) with a preconditioner given by $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$; (in figures denoted by "PCG");

- MF solver, Eq. (32); (in figures denoted by "MF");

- MF method within three different cooling schemes as proposed in [4], [9], and [16], respectively. In the first one, the value of the cooling parameter λ is adjusted adaptively after each iteration. This scheme requires an a priori knowledge (estimate) on the error measure (see Eq. (33) below) of the solution. For the purpose of the experiments, this is tightly approximated using the solution of the PCG solver. The scheme of [9] defines a discrete grid of logarithmically spaced values of λ, which spans the range from 1 up to some suitable maximal value, which in our runs we set to $\lambda_{max} = 10^4$. For each value of λ, a fixed number of iterations is performed. Hereafter, we adopt the scheme 16×10, i.e., we perform 10 iterations for each of the 16 values of λ. For the case with partial sky observations, we continue with the fixed-point iterations (32) for $\lambda = 1$. The cooling scheme of [16, Algorithm 1] reduces λ by a constant factor, η, so $\lambda \rightarrow \lambda \times \eta$ and iterates as long as two consecutive approximations satisfy $\|\mathbf{s}^{(i)} - \mathbf{s}^{(i-1)}\|/\|\mathbf{s}^{(i)}\| < \epsilon$. In our experiments we start with $\lambda_{max} = 10^4$, we set $\eta = 3/4$ and $\epsilon \equiv 10^{-4}$.

We start the iterations with a vector of zeros as an initial guess.

The signal covariance, $\mathbf{S}$, is computed assuming the CMB power spectra as used for the simulations. The noise covariance is block-diagonal in the pixel domain but not proportional to $\mathbf{I}$ as the noise is assumed to be inhomogeneous. It is taken to be exactly the same as the noise covariance used for the simulations.

For all solvers the inverse signal covariance is applied to a map-length pixel domain vector in the harmonic domain, when first the vector is represented by a vector of its harmonic coefficients computed with help of a HEALpy routine, `map2alm`, these are subsequently weighted by the inverse of the power spectra and transformed back to the pixel domain using HEALpy's `alm2map` routine. The inverse noise covariance is applied to any pixel-domain vector directly with the elements corresponding to unobserved pixels set to 0.

We note that we always estimate the Wiener-filtered sky signal over the full sky. In all cases shown below we apply all the solvers to exactly the same input data sets.

We have validated our implementations by considering a simplified data set with white noise. In this case, the noise covariance $\mathbf{N}$ is proportional to the unit matrix, and the PCG solver with preconditioner $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$, and the MF solver converged to within the numerical precision in a single step as expected, as in this case $\mathbf{T} = \mathbf{N}$. Moreover, in the cases of the actual simulated data sets used in our test, the results obtained with the different solvers are consistent.

4.3.1 Convergence metric.

The Wiener Filter problem (26) can be recast as a minimization of the functional,

$$\chi^2(\mathbf{x}) = \mathbf{x}^t \mathbf{S}^{-1} \mathbf{x} + (\mathbf{m} - \mathbf{x})^t \mathbf{N}^{-1} (\mathbf{m} - \mathbf{x}). \tag{33}$$

---

[2]We use specifically the noise pattern for the 9-year observation of the V-band, available from `https://lambda.gsfc.nasa.gov/product/map/dr5/maps_band_r9_iqu_9yr_get.cfm`.

[3]`http://irsa.ipac.caltech.edu/data/Planck/release_2/ ancillary-data/previews/HFI_Mask_GalPlane-apo2_2048_R2.00/index.html`

Indeed, we have

$$\arg\min \chi^2(\mathbf{x}) = \left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)^{-1} \mathbf{N}^{-1}\mathbf{m} = \mathbf{s}_{WF}.$$

Algebraic manipulations show that χ^2 is directly related to the energy norm, as we have

$$\|\mathbf{s}_{WF} - \mathbf{x}\|_{\mathbf{S}^{-1}+\mathbf{N}^{-1}}^2 = \chi^2(\mathbf{x}) - \chi^2(\mathbf{s}_{WF}).$$

We can thus use physically motivated χ^2 as a convergence measure instead of the energy norm. We note that we expect that

$$\chi^2(\mathbf{x}) \to \chi^2(\mathbf{s}_{WF}) = \mathbf{m}^t (\mathbf{S} + \mathbf{N})^{-1} \mathbf{m}. \tag{34}$$

As with the energy norm, Sect. 3.1, this asymptotic value of χ^2 can not be straightforwardly computed without knowing the Wiener filter estimate precisely. However, we expect that it should be close to $\langle\chi^2(\mathbf{s}_{WF})\rangle = n_{Stokes} n_{pix} \equiv n_{DOF}$ within a small scatter on order of $O(\sqrt{2 n_{Stokes} n_{pix}})$, if our assumptions about the sky signal and the noise are correct. Here the angle brackets denote the average over an ensemble of sky and noise realizations and n_{Stokes} stands for the number of considered Stokes parameters and is equal to 3 for most of our tests. We can therefore define the convergence in this case by requiring that the incremental change of χ^2 between consecutive iterations is not larger than some small fraction of $\langle\chi^2(\mathbf{s}_{WF})\rangle$ [4]. If the absolute value of the final χ^2 is statistically inconsistent with the expected one, this could be an indication of prematurely stalled convergence or of a problem with the model assumed for the measured data, $\mathbf{m}$.

Given the discussion of Sect. 3.1, we expect that in terms of minimizing the χ^2-measure, the PCG method with preconditioner $\mathbf{M} = \mathbf{S}^{-1} + \mathbf{T}^{-1}$ should be superior to the fixed-point iterations, Eq. (32).

In addition to the χ-measure we also plot the norm of the residual corresponding to the (preconditioned) problem as suggested in [16]. This is given by,

$$\mathbf{S}^{1/2}\left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)\mathbf{S}^{1/2}\mathbf{S}^{-1/2}\mathbf{s}_{WF} = \mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m}; \tag{35}$$

[16, Appendix C]. This system is significantly better conditioned than the original one, Eq. (26). The corresponding relative norm of the residual reads then,

$$\frac{\|\mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m} - \mathbf{S}^{1/2}\left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)\mathbf{x}\|}{\|\mathbf{S}^{1/2}\mathbf{N}^{-1}\mathbf{m}\|} = \frac{\|\mathbf{N}^{-1}\mathbf{m} - \left(\mathbf{S}^{-1} + \mathbf{N}^{-1}\right)\mathbf{x}\|_{\mathbf{S}}}{\|\mathbf{N}^{-1}\mathbf{m}\|_{\mathbf{S}}}. \tag{36}$$

4.3.2 Performance.

Figure 1 shows a comparison between the PCG, Eq. (31), and the MF, Eq. (32), solvers as applied to the Wiener Filter problem. As expected PCG indeed reduces the error significantly faster.

In Figs. 2 and 3, we compare PCG (with the MF preconditioner $\mathbf{S}^{-1} + \mathbf{T}^{-1}$ and $\lambda = 1$) with the MF method using the adaptive cooling schemes described above. We can see that PCG yields robust performance in all the test cases. In the case with full sky observations, the MF solvers (with or without cooling) reach on the level 10^{-7} of the relative $\mathbf{S}$-norm of the residual, their asymptotic convergence rate and exhibit a plateau of convergence. This is not the case of PCG solver that converges to the machine precision level. In the experiment with partial sky coverage, we observe a decrease of the convergence rate for the PCG as well as the MF solvers due to significantly worse conditioning of the problem. However, even in this case, the PCG method is superior to the MF solvers. We expect that using more advanced preconditioners, which can alleviate the effect of very small eigenvalues can bring a further significant improvement.

We, however, note that PCG appears to be outperformed by the MF method with the [4] cooling proposal within the first ten or so iterations. As both methods solve then different linear systems, due to different values of λ, this does not contradict our conclusions in Sect. 3.1. This also does not change the overall assessment of the relative merit of both these techniques as no convergence is then ever reached in terms of any of the considered metrics. We discuss a possible origin of this behavior later on, Sect. 5.3, in the map-making context, where we suggest a simple antidote for it that could potentially further ameliorate performance of the PCG approach.
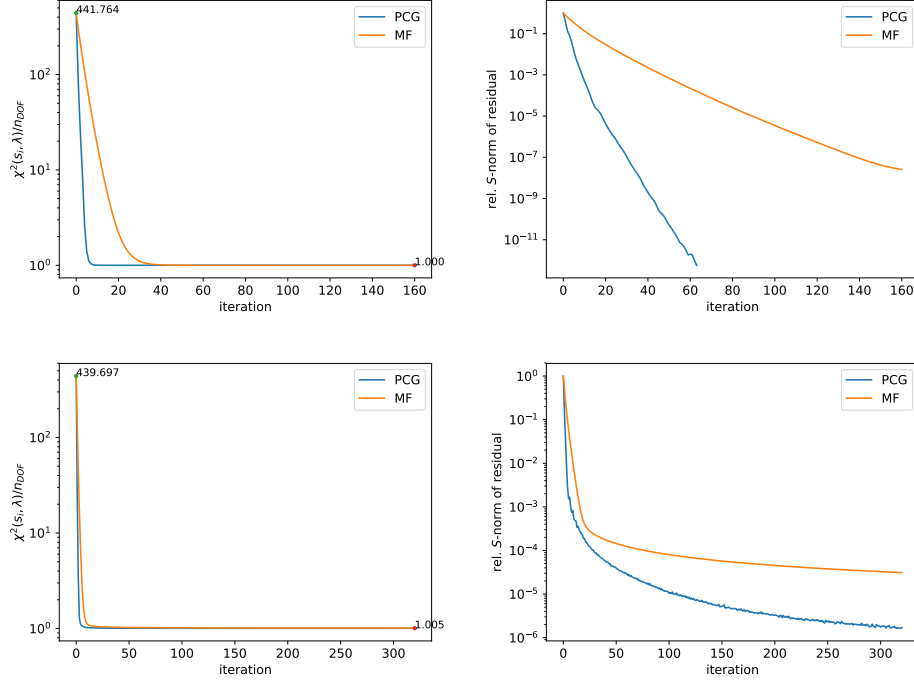
Figure 1: Convergence of PCG and MF methods using two different convergence measure: χ^2, Eq. (33), left panels, and the **S**-weighted relative residual, Eq. (36), right panels. The top/bottom row shows the case with the full/partial sky coverage.
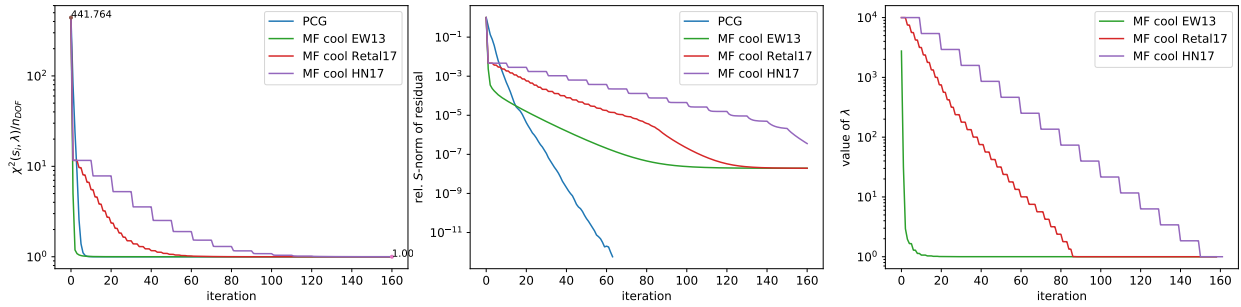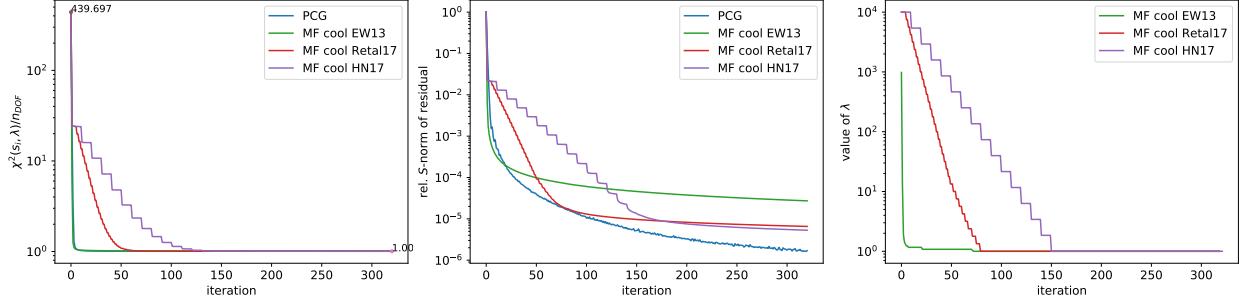

Figure 2: Comparison of the convergence of the PCG solver and the MF technique with an adaptive cooling for different cooling prescriptions and using different convergence criteria: the χ^2, left, the **S**-weighted relative residual, middle. The right panel shows the values of λ as a function of the iteration as adapted by the different cooling schemes. These results are for the data sets with the full sky coverage.

5 Application to map-making.

5.1 The problem.

Data collected by modern, single-dish CMB experiments are modeled as

$$\mathbf{d} = \mathbf{P}\,\mathbf{m} + \mathbf{n}, \tag{37}$$

Figure 3: As Fig. 2 but for the data set with the partial sky coverage, $f_{sky} = 0.2$.

where $\mathbf{d}$ stands for a vector of all measurements, $\mathbf{m}$ is a pixelized map of the sky signal and $\mathbf{n}$ is the instrumental noise. We assume for simplicity that experimental beams are axially symmetric and that the sky signal $\mathbf{m}$ is already convolved with the beam. In this case, pointing matrix $\mathbf{P}$ simply defines which pixel of the map, $\mathbf{m}$, is observed at each measurement and with what weight it contributes to the measurement. In such case, the pointing matrix is very sparse as it contains only one, for the total intensity measurements, or three, for the polarization-sensitive ones, non-zero elements per each row. Moreover, $\mathbf{P}^t \mathbf{P}$ is either diagonal or block-diagonal with 3×3 blocks. If we assume that the instrumental noise is Gaussian with the covariance given by $\mathbf{N}$, a maximum-likelihood estimate of the sky signal can be written as

$$\mathbf{m}_{ML} \equiv \left(\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}\right)^{-1} \mathbf{P}^t \mathbf{N}^{-1} \mathbf{d}, \tag{38}$$

and therefore requires a solution of large linear system. The sizes of the involved object vary wildly depending on the experiment but the number of pixels in the map $\mathbf{m}$ can easily reach $O(10^6)$, while the number of measurements, $O(10^{12-15})$. Consequently, the system could be only solved iteratively explicitly capitalizing on the structure and sparsity of the involved data objects.

Traditionally (e.g. [3, 2]) the iterative method of choice was a PCG technique with a simple preconditioner given by

$$\mathbf{M} = \mathbf{P}^t \operatorname{diag}(\mathbf{N}^{-1}) \mathbf{P}. \tag{39}$$

Hereafter we refer to this standard preconditioner as block-diagonal or Jacobi. However, more involved preconditioners have been also considered and found successful, e.g., [6, 21, 12, 15].

More recently, [9] (see also [8]) have proposed to apply the messenger-field technique to the map-making problem. Below, we discuss the approach of this former work. The proposal here is again to split the noise covariance into two parts, $\mathbf{N} \equiv \bar{\mathbf{N}} + \mathbf{T}$, where $\mathbf{T} = \tau \mathbf{I}$, $\tau = \min(\operatorname{diag}(\mathbf{N}))$, using then Eq. (27) we can rewrite the system matrix of the map-making equation, Eq. (38), as

$$\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P} = \mathbf{P}^t \mathbf{T}^{-1} \mathbf{P} - \mathbf{P}^t \mathbf{T}^{-1} \left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1} \mathbf{T}^{-1} \mathbf{P}, \tag{40}$$

Where the first term on the right hand side corresponds to matrix $\mathbf{C}$ and the second one to matrix $\mathbf{D}$ as defined in Eq. (4). Following the formalism from Sect. 2.1 we can now write the messenger-field equations for this system, which read

$$\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)\mathbf{t} = \bar{\mathbf{N}}^{-1} \mathbf{d} + \mathbf{T}^{-1} \mathbf{P} \mathbf{m}_{ML}, \tag{41}$$

$$\left(\mathbf{P}^t \mathbf{T}^{-1} \mathbf{P}\right)\mathbf{m}_{ML} = \mathbf{P}^t \mathbf{T}^{-1} \mathbf{t}, \tag{42}$$

with the messenger field $\mathbf{t}$ appearing explicitly, or

$$\left[\mathbf{I} - \left(\mathbf{P}^t \mathbf{T}^{-1} \mathbf{P}\right)^{-1} \mathbf{P}^t \mathbf{T}^{-1} \left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1} \mathbf{T}^{-1} \mathbf{P}\right]\mathbf{m}_{ML} = \left(\mathbf{P}^t \mathbf{T}^{-1} \mathbf{P}\right)^{-1} \mathbf{P}^t \mathbf{N}^{-1} \mathbf{d}, \tag{43}$$

without it. (N.B., note that unlike in Eq. (6) the matrix $\mathbf{P}^t$ is taken out of the definition of the messenger field.) Solving any of these two sets of equations using fixed-point iterations is equivalent to the messenger-field solver. For comparison we will also solve the last equation using CG technique. The latter is equivalent to solving the map-making equation, Eq. (38), using a PCG method with the preconditioner taken to be $\mathbf{M} = \mathbf{P}^t \mathbf{T}^{-1} \mathbf{P}$, which in the case under consideration is equivalent to the standard preconditioner.

5.2 Simulated data.

We simulate mock time ordered data $\mathbf{d}$ as a sum of two terms, one corresponding to the sky signal and the other to instrumental noise. These are computed as,

$$\mathbf{d}_t = \mathbf{I}_{p(t)} + \mathbf{Q}_{p(t)} \cos 2\varphi(t) + \mathbf{U}_{p(t)} \sin 2\varphi(t) + \mathbf{n}_t, \tag{44}$$

where $p(t)$ denotes the sky pixels observed at time t and $\varphi(t)$ – the corresponding orientation of the polarizer. The signal terms are read off from the signal only maps of Stokes parameters I, Q, and U, following assumed scanning strategy defined by $p(t)$ and $\varphi(t)$. These maps are produced in the Healpix pixelization with the resolution parameter n_{side} set to 1024. These signals are random realizations of the CMB anisotropies corresponding to the currently preferred cosmological model [13].

   We produce two data sets with different statistical properties. Each data set comprises the data of a single detector that are however scaled to represent the performance of an entire detector array and thus being more representative of the current data. In both cases we assume a raster scan pattern in the sky coordinates, when a rectangular sky patch made of 256×256 Healpix pixels is scanned either horizontally, i.e., in right ascension or vertically, i.e., in declination. The patch is centered at the equator. The length of the simulated data vector is the same for the two cases and roughly equal to 10^8.

   In the first data set, the sky patch is first scanned horizontally and later vertically. The horizontal scanning assumes 256 complete, i.e., left-to-right followed by right-to-left, sweeps when each pixel is sampled on average 4 times on each sweep. Once this is done the declination is changed and the new horizontal scan commences. This is repeated 256 times with each horizontal scan corresponding to a different row of the Healpix pixels. The vertical scan is implemented in a similar way.

   In this case we assume that the polarizer direction is quickly modulated so the full 2π angle is sampled within each single crossing of each sky pixel. This is ensured by setting the polarizer angle in the sky coordinates to follow a repeating sequence of $0, \pi/4, \pi/2, 3\pi/4$. In practice, this could mimic the case of an experiment using a smoothly rotating half wave plate.

   For the second data set, we divide it into 4 equal consecutive subsets, each of which implements the same raster scan made of horizontal scanning within the first half of the subset followed by the vertical scan in the second half. However, the scanning is assumed to be faster and there is only 1 sample taken per pixel for each pixel crossing. This ensures the same data length. For each subsets the angle of the polarizer in the sky coordinates is fixed and equal to $0, \pi/4, \pi/2, 3\pi/4$. This scanning strategy mimics an experiment where the polarizer is stepped discretely only after each of the four subscans.

   We simulate the instrumental noise as a correlated noise with a power spectrum given by

$$\mathcal{P}(f) \equiv \sigma^2 \left(1 + \frac{f_{knee}}{f}\right), \tag{45}$$

where $1/f_{knee}$ is taken to be ~ 500 times longer than the sampling rate, i.e., corresponding to a length of a single full sweep of the sky. We further apodize the low frequency noise effectively flattening the noise spectrum for frequencies lower than a tenth of the knee frequency.

   We consider two different noise levels, one ensuring rather high signal to noise of the resulting maps, with rms noise $\sigma_{rms}^{Q,U} = \sigma_{rms}^I \cdot \sqrt{2} \simeq 2\mu K$ for the recovered Q and U maps and the other with lower signal-to-noise corresponding to $\sigma_{rms}^{Q,U} \simeq 30\mu K$. We refer to these cases as the low and high noise data.

   We note, that if the instrumental noise were white then both scanning strategies would have been equivalent and the standard, block-diagonal preconditioners in both these cases would have been identical. This is however not the case in the presence of the correlated noise. In fact we expect that the off-diagonal noise correlation of the recovered

Q and U maps should be small for the first data set with the quickly rotating HWP while they should be non-negligible in the case of the second data set with the stepped polarizer, leading to a different convergence patterns of the studied solvers.

5.3 Numerical results.

5.3.1 Convergence metric.

For measuring the error of an approximation $\mathbf{x}$ we consider, following [9], a χ-measure

$$\chi^2(\mathbf{x}) = (\mathbf{d} - \mathbf{Px})^t \mathbf{N}^{-1} (\mathbf{d} - \mathbf{Px}). \tag{46}$$

Analogously to the Wiener Filtering application, this measure is minimized by the maximum-likelihood estimate (38) and it is equivalent to the energy norm of the error, Eq. (15), with respect to the system matrix, $\mathbf{A} \equiv \mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}$. Indeed,

$$\chi^2(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_{ML}\|_{\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}} + + \mathbf{d}^t (\mathbf{N}^{-1} - \mathbf{N}^{-1} \mathbf{P} (\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^t \mathbf{N}^{-1}) \mathbf{d}, \tag{47}$$

and thus,

$$\chi^2(\mathbf{m}_{ML}) = \mathbf{d}^t (\mathbf{N}^{-1} - \mathbf{N}^{-1} \mathbf{P} (\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^t \mathbf{N}^{-1}) \mathbf{d}. \tag{48}$$

As before this value is not directly available. However, we can compute the average value of $\chi^2(\mathbf{m}_{ML})$ over the statistical ensemble of the input data realizations and use it as a benchmark for the convergence using the χ^2-measure. This can be done analytically on observing that the matrix on the right hand side of Eq. (48) is a projection operator, which projects out all time-domain modes, which are sky stationary, i.e., objects of the form $\mathbf{P}\mathbf{y}$ for some arbitrary pixel-domain object $\mathbf{y}$. If so,

$$\chi^2(\mathbf{m}_{ML}) = \mathbf{n}^t (\mathbf{N}^{-1} - \mathbf{N}^{-1} \mathbf{P} (\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^t \mathbf{N}^{-1}) \mathbf{n}, \tag{49}$$

and

$$\begin{aligned} \langle \chi^2(\mathbf{m}_{ML}) \rangle_{noise} &= \operatorname{tr} (\mathbf{I} - \mathbf{N}^{-1/2} \mathbf{P} (\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^t \mathbf{N}^{-1/2}) \\ &= n_t - n_{Stokes} n_{pix} \equiv n_{DOF}, \end{aligned} \tag{50}$$

where n_t and n_{pix} denote the sizes of the data set in time and pixel domains respectively and we have assumed that the system matrix, $\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}$, is non-singular, and used the fact that $\mathbf{N}^{-1/2} \langle \mathbf{n}\,\mathbf{n}^t \rangle \mathbf{N}^{-1/2} = \mathbf{I}$. Having this value, we then define the convergence criterion in terms of the χ^2-measure by requiring that the incremental change of χ^2 between two consecutive iterations is sufficiently small as compared to n_{DOF}.

We note that in the figures in order to make the behavior of the χ-measure more conspicuous, instead of the χ^2 itself, we plot its relative difference with respect to the minimal value of χ^2 derived within the PCG iterations, which we denote as $\chi^2(\mathbf{m}_{min})$, and the plotted quantity is then given by,

$$\chi^2(\mathbf{x})/\chi^2(\mathbf{m}_{min}) - 1. \tag{51}$$

For completeness we also plot the standard relative residual defined as,

$$\frac{\|\mathbf{P}^t \mathbf{N}^{-1} \mathbf{d} - (\mathbf{P}^t \mathbf{N}^{-1} \mathbf{P}) \mathbf{x}\|}{\|\mathbf{P}^t \mathbf{N}^{-1} \mathbf{d}\|}. \tag{52}$$

5.3.2 Performance.

We reconstruct the sky signal from the data using different solvers as discussed here and compare their relative performance. In all cases, we use the same pixelization, i.e., Healpix with $n_{side} = 1024$, for the recovered maps as we used for the simulations. We have validated our implementation by running cases with noise-free data and recovering the
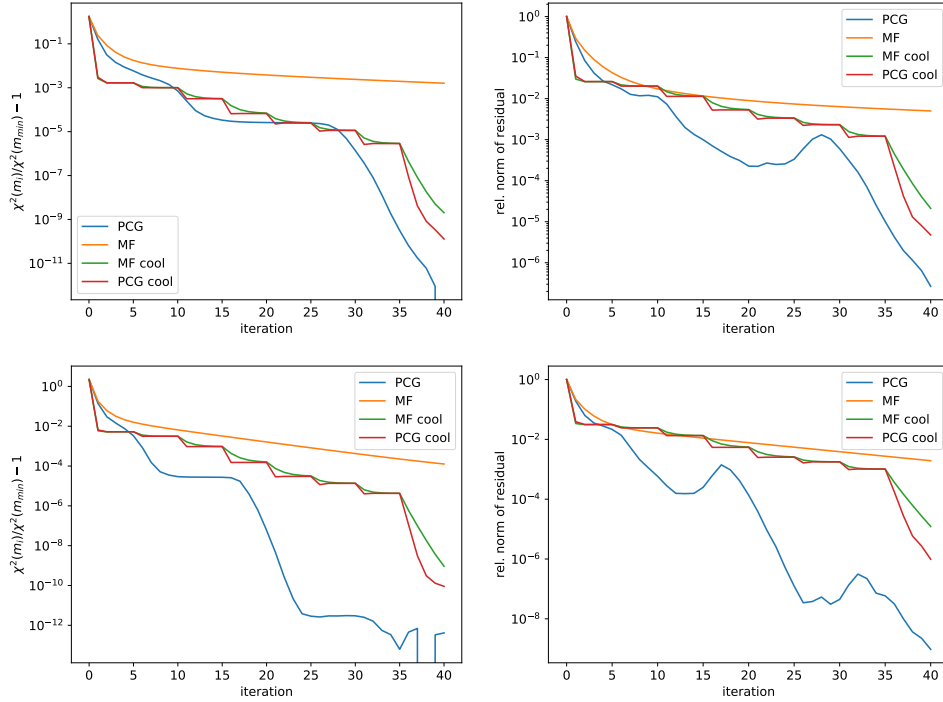
Figure 4: Comparison of the convergence for the PCG, the messenger-field methods standalone and incorporated within a cooling scheme, 8×5, for the 1st, top row, and the 2nd, bottom row, simulated data set and assuming the low noise level.

input maps within the numerical precision. We have also found that the results produced by different solvers for each of the data sets agree.

We first compare the convergence of the PCG solver, of the MF iterations without cooling, and of the PCG and MF with the 8×5 cooling scheme. The results for the low noise cases are given in Figs. 4 and the high noise cases are shown in Fig. 5.

We can see that, as claimed in [9], the MF with cooling technique indeed reaches higher accuracy in comparison to MF without cooling. However, the standard PCG is in these experiments still superior. As in the Wiener Filter application, Sect. 4.3, we observe that in the low noise cases the cooling technique, used either with PCG or with MF, improves more rapidly on the solution within the first iterations than the PCG method with no cooling. We attribute this to the fact that during those initial iterations the cooling method solves a modified system of the initial equations with an assumed large value of λ. The approximate solution derived on these first iteration is then equivalent to a simple binned map. This for the low noise cases provides a good rendition of the sky signal, thus leading to an abrupt decrease of the residuals. In the absence of the cooling the PCG technique initiated with vector of zeros needs to perform at least a few iterations to reach a comparably good solution. We can however improve on the performance of the standalone PCG by using a simple binned map given by

$$\mathbf{m}^{(0)} = (\mathbf{P}^t \operatorname{diag}(\mathbf{N})^{-1} \mathbf{P})^{-1} \mathbf{P}^t \operatorname{diag}(\mathbf{N})^{-1} \mathbf{d}, \tag{53}$$

as the starting vector for the PCG solver. Such a map is quickly computable and thus can be readily available at the onset of the solution.

We illustrate these considerations in Fig. 6, where we compare the convergence of the PCG run with the initial vector made of zeros and the convergence of the solvers: standalone PCG, MF with cooling, and PCG with cooling, assuming $\mathbf{m}^{(0)}$, Eq. (53), as the initial vector. The results shown are for the low noise case. As expected there is a significant improvement in the overall performance of the PCG method relatively to the other solvers but also as
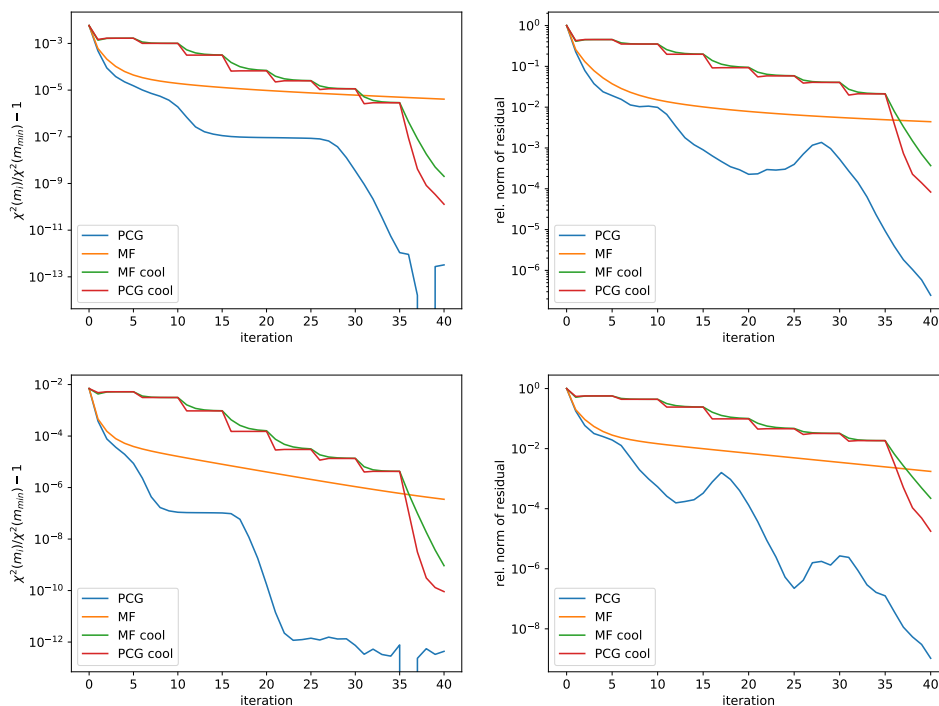
Figure 5: As Fig. 4 but for the high noise level.

compared to the case of the vanishing initial guess. This showcases the importance of the appropriate choice of the initial guess for the PCG approaches in the cases of high signal-to-noise solution. As the new CMB data sets target predominantly CMB B-mode polarization, the maps of Stokes parameters will be increasingly having very high signal-to-noise and this observation may be therefore of importance for their data analysis.
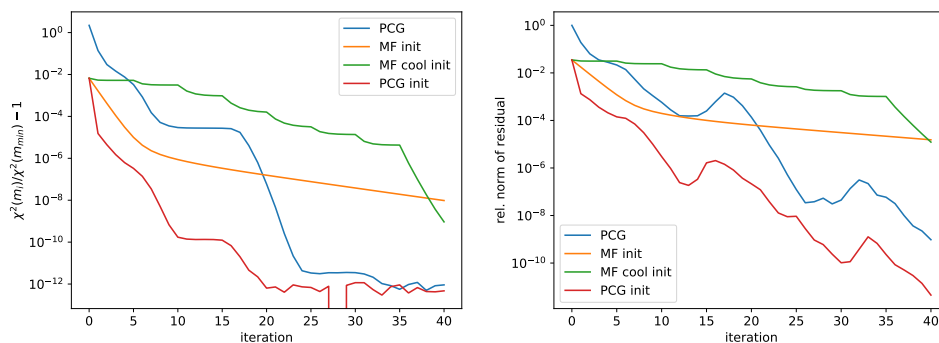

Figure 6: Comparison of the convergence rates of different iterative solvers for a non-zero starting vector, $\mathbf{m}^{(0)}$, as given in Eq. (53). For comparison the blue curve shows the case of the PCG with a starting vector of zeros. The results are for the 2nd scanning strategy and the low noise case.

6 Conclusions.

We have shown that the messenger-field solvers of sets of linear equations perform fixed-point iterations of an appropriately preconditioned system of equations. Consequently, in general they are expected to display inferior performance to that of a conjugate gradient solver applied to the same preconditioned systems or, equivalently, to that of a PCG solver with the same preconditioner as implicitly used in the messenger-field method in order to precondition the initial problem. We have backed up this contention with analytic arguments and demonstrated it on numerical experiments involving two applications drawn from modern CMB data analysis practice: Wiener Filters and CMB map-making problem. We have compared the performance of the studied methods from the perspective of the number of iterations needed to reach convergence. However, our conclusions are expected to be directly applicable also to considerations involving time-to-solution, as the computational cost per iteration incurred in the different methods is found to be roughly comparable.

We thus conclude that looking towards the future advanced preconditioning coupled with the conjugate gradient technique offers more promise for more expeditious solvers that the current ones than the messenger-field approach does. While at this time, the PCG solvers with the standard, block-diagonal, Eq. (39), preconditioner in the map-making case, and the preconditioner given by $\mathbf{S}^{-1} + \mathbf{T}^{-1}$, Eq. (28), in the Wiener filter case, with a potentially appropriately adapted initial guess should outperform the currently proposed messenger-field approaches. We also note that better preconditioners have been already proposed in particular in the map-making context e.g. [6, 21]. This notwithstanding, the messenger-field approach may still be found of interest in some specific applications.

In the context of the PCG methods, we have found that the convergence may be sped up by an appropriate choice of the initial vector. While the gain is largely negligible for the cases with a low signal-to-noise solution, it can become significant if the solution is expected to have high signal content. We have found this effect particularly relevant for the map-making procedure, where we have showed that the choice of the simple binned map as the initial vector can result in a significant improvement of the map-making solver convergence.

A Implementation of PCG for Wiener Filter.

In the context of solving the Winer Filter problem, Eq. (25), each step of fixed-point method, Eq. (12), resp. Eq. (32), requires one direct and one inverse spherical harmonic transforms that are assumed to be the most time consuming elements of the solution process. To keep the same number of transforms in each PCG iteration, we cannot apply first matrix $\mathbf{A} = \mathbf{S}^{-1} + \mathbf{N}^{-1}$ and then precondition the residual by $\mathbf{C}^{-1} \equiv (\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}$ as it is done in one of the standard PCG implementations listed below in Algorithm A.1. This implementation involves two direct and two inverse transforms: in the evaluation of $\mathbf{A}\mathbf{p}^{(i-1)}$ and in $\mathbf{C}^{-1}\mathbf{r}^{(i)}$.

Using the formula for $\mathbf{p}^{(i)}$, $\widetilde{\mathbf{r}}^{(i)}$ and the form of the matrix $\mathbf{A} = \mathbf{C} - \mathbf{D}$, we can write

$$\mathbf{A}\mathbf{p}^{(i)} = \mathbf{A}\big(\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{p}^{(i-1)}\big) = \mathbf{C}\mathbf{C}^{-1}\mathbf{r}^{(i)} - \mathbf{D}\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)} = \mathbf{r}^{(i)} - \mathbf{D}\widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)}.$$

Therefore the vector $\mathbf{A}\mathbf{p}^{(i)}$ can be computed recursively without spherical harmonic transforms and the cost of one PCG iteration is the same (in terms of spherical harmonic transforms) as the cost of one iteration of the fixed-point method, Eq. (12).

Another formula that proved in our numerical experiments to be more stable (yet slightly more costly) is to evaluate at once the vectors $\mathbf{C}^{-1}\mathbf{r}^{(i)}$ and $\mathbf{S}^{-1}\mathbf{C}^{-1}\mathbf{r}^{(i)}$ (recall that $\mathbf{A}\mathbf{C}^{-1}\mathbf{r}^{(i)} = (\mathbf{S}^{-1} + \mathbf{N}^{-1})\mathbf{C}^{-1}\mathbf{r}^{(i)}$). This can be done using direct spherical harmonic transform of one vector, $\mathbf{r}^{(i)}$, and inverse spherical harmonic transform of two vectors[4]. Then we simply update $\mathbf{A}\mathbf{p}^{(i)} = \mathbf{A}\mathbf{C}^{-1}\mathbf{r}^{(i)} + \delta^{(i)}\mathbf{A}\mathbf{p}^{(i-1)}$.

---

[4]with the typical computational cost significantly smaller than twice the cost of one single inverse spherical harmonic transform.

Algorithm A.1 PCG for $\mathbf{As} = \mathbf{b}$ with the preconditioner $\mathbf{C}$

Given $\mathbf{s}^{(0)}$, $\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{As}^{(0)}$, $\widetilde{\mathbf{r}}^{(0)} = \mathbf{C}^{-1}\mathbf{r}^{(0)}$, $\mathbf{p}^{(0)} = \widetilde{\mathbf{r}}^{(0)}$.

For $i = 1, 2, \ldots$

$$\gamma^{(i-1)} = \frac{(\mathbf{r}^{(i-1)}, \widetilde{\mathbf{r}}^{(i-1)})}{(\mathbf{p}^{(i-1)}, \mathbf{Ap}^{(i-1)})},$$

$$\mathbf{s}^{(i)} = \mathbf{s}^{(i-1)} + \gamma^{(i-1)}\mathbf{p}^{(i-1)},$$

$$\mathbf{r}^{(i)} = \mathbf{r}^{(i-1)} - \gamma^{(i-1)}\mathbf{Ap}^{(i-1)},$$

$$\widetilde{\mathbf{r}}^{(i)} = \mathbf{C}^{-1}\mathbf{r}^{(i)},$$

$$\delta^{(i)} = \frac{(\mathbf{r}^{(i)}, \widetilde{\mathbf{r}}^{(i)})}{(\mathbf{r}^{(i-1)}, \widetilde{\mathbf{r}}^{(i-1)})},$$

$$\mathbf{p}^{(i)} = \widetilde{\mathbf{r}}^{(i)} + \delta^{(i)}\mathbf{p}^{(i-1)}.$$

Moreover, the properties of (P)CG also allow to evaluate the decrease of the χ-measure without computing it explicitly using Eq. (33) (this computation involves $\mathbf{S}^{-1}$ and therefore also direct and inverse spherical harmonic transforms). The evaluation proposed below is numerically stable; see a thorough analysis in [20]. There holds

$$\|\mathbf{s} - \mathbf{s}^{(0)}\|_{\mathbf{A}}^2 = \sum_{j=1}^{i} \gamma_j \, (\mathbf{r}^{(j)}, \widetilde{\mathbf{r}}^{(j)}) + \|\mathbf{s} - \mathbf{s}^{(i)}\|_{\mathbf{A}}^2.$$

Using the above discussion on the relationship between the energy norm and the χ-measure, we have

$$\chi^2(\mathbf{s}^{(0)}) = \sum_{j=1}^{i} \gamma_j \, (\mathbf{r}^{(j)}, \widetilde{\mathbf{r}}^{(j)}) + \chi^2(\mathbf{s}^{(i)}).$$

After computing $\chi^2(\mathbf{s}^{(0)})$ (that is for zero initial approximation $\mathbf{s}^{(0)} = 0$ equal to $\mathbf{m}^t\mathbf{N}^{-1}\mathbf{m}$) we can therefore simply evaluate the χ-measure in every PCG iteration using already computed scalar quantities without any additional spherical harmonic transforms.

B   Proof of convergence of the messenger-field method.

In this appendix we prove that the messenger-field method for Wiener Filter is (asymptotically) converging. An analogous proof can be used also for the map-making application.

Following the discussion below the Eq. (13), we prove the convergence by showing that the eigenvalues of $\mathbf{C}^{-1}\mathbf{D}$ are in the absolute value smaller than unity. We start by rewriting Eq. (30) as the system

$$\underbrace{\begin{pmatrix} \bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1} & -\mathbf{T}^{-1} \\ -\mathbf{T}^{-1} & \mathbf{S}^{-1} + \mathbf{T}^{-1} \end{pmatrix}}_{\equiv \mathbb{A}} \begin{pmatrix} \mathbf{t} \\ \mathbf{s}_{WF} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{N}}^{-1}\mathbf{m} \\ 0 \end{pmatrix}$$

The reduced system (after the elimination of the messenger field $\mathbf{t}$) then corresponds to forming the Schur complement $\mathbb{S}$ of $\mathbb{A}$,

$$\mathbb{S} \equiv \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right) - \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}$$

and solving

$$\mathbb{S}\mathbf{s}_{WF} = \mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\bar{\mathbf{N}}^{-1}\mathbf{m}.$$

The MF iterations are obtained by multiplying (preconditioning) the above system by $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}$ from the left.

Now we show the bounds on the eigenvalues of

$$\mathbf{C}^{-1}\mathbf{D} = \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbf{T}^{-1} = \mathbf{I} - \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbb{S}.$$

Since $\mathbb{A}$ is an SPD matrix, its Schur complement $\mathbb{S}$ is also SPD. Moreover, the spectrum of $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}$ satisfies

$$\Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbb{S}\right) = \Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\mathbb{S}\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\right)$$

and therefore the eigenvalues of $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1}\mathbb{S}$ are positive. Plugging into the above equation the formula for the Schur complement $\mathbb{S}$, we have

$$\Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbb{S}\right) = \Lambda\left(\mathbf{I} - \left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\right)$$
$$= 1 - \Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\left(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1}\right)^{-1}\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1/2}\right).$$

The matrix $(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}(\bar{\mathbf{N}}^{-1} + \mathbf{T}^{-1})^{-1}(\mathbf{S}^{-1} + \mathbf{T}^{-1})^{-1/2}$ is SPD. Altogether,

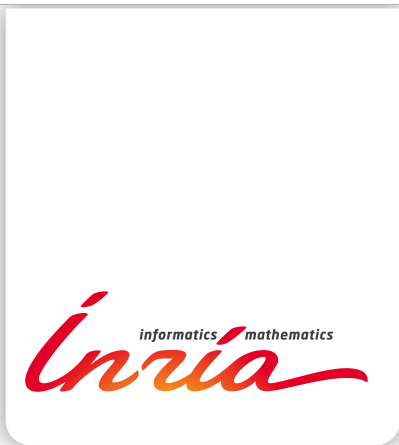$$\Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbb{S}\right) \in (0, 1).$$

Consequently,

$$\Lambda\left(\mathbf{C}^{-1}\mathbf{D}\right) = 1 - \Lambda\left(\left(\mathbf{S}^{-1} + \mathbf{T}^{-1}\right)^{-1}\mathbb{S}\right) \in (0, 1),$$

which proves that the MF iterations are asymptotically converging.

References

[1] E. F. Bunn, K. B. Fisher, Y. Hoffman, O. Lahav, J. Silk, and S. Zaroubi. Wiener filtering of the COBE Differential Microwave Radiometer data. *ApJ*, 432:L75–L78, Sept. 1994.

[2] C. M. Cantalupo, J. D. Borrill, A. H. Jaffe, T. S. Kisner, and R. Stompor. MADmap: A Massively Parallel Maximum Likelihood Cosmic Microwave Background Map-maker. *ApJS*, 187:212–227, Mar. 2010.

[3] G. de Gasperis, A. Balbi, P. Cabella, P. Natoli, and N. Vittorio. ROMA: A map-making algorithm for polarised CMB data sets. *A&A*, 436:1159–1165, June 2005.

[4] F. Elsner and B. D. Wandelt. Efficient wiener filtering without preconditioning. *A&A*, 549:A111, 2013.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[6] L. Grigori, R. Stompor, and M. Szydlarski. A parallel two-level preconditioner for cosmic microwave background map-making. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC '12, pages 91:1–91:10, Los Alamitos, CA, USA, 2012. IEEE Computer Society Press.

[7] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, PA, second edition, 2002.

[8] K. M. Huffenberger. Preconditioner-free Wiener filtering with a dense noise matrix. *MNRAS*, Feb. 2018.

[9] K. M. Huffenberger and S. K. Næss. Cosmic Microwave Background Mapmaking with a Messenger Field. *ApJ*, 852:92, Jan. 2018.

[10] M. A. Janssen and S. Gulkis. Mapping the sky with the COBE differential microwave radiometers. In M. Signore and C. Dupraz, editors, *NATO Advanced Science Institutes (ASI) Series C*, volume 359 of *NATO Advanced Science Institutes (ASI) Series C*, pages 391–408, 1992.

[11] J. Liesen and Z. Strakoš. *Krylov Subspace Methods: Principles and Analysis*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[12] S. K. Næss and T. Louis. A fast map-making preconditioner for regular scanning patterns. *J. Cosmology Astropart. Phys.*, 8:045, Aug. 2014.

[13] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, Sept. 2016.

[14] D. Poletti, G. Fabbian, M. Le Jeune, J. Peloton, K. Arnold, C. Baccigalupi, D. Barron, S. Beckman, J. Borrill, S. Chapman, Y. Chinone, A. Cukierman, A. Ducout, T. Elleflot, J. Errard, S. Feeney, N. Goeckner-Wald, J. Groh, G. Hall, M. Hasegawa, M. Hazumi, C. Hill, L. Howe, Y. Inoue, A. H. Jaffe, O. Jeong, N. Katayama, B. Keating, R. Keskitalo, T. Kisner, A. Kusaka, A. T. Lee, D. Leon, E. Linder, L. Lowry, F. Matsuda, M. Navaroli, H. Paar, G. Puglisi, C. L. Reichardt, C. Ross, P. Siritanasak, N. Stebor, B. Steinbach, R. Stompor, A. Suzuki, O. Tajima, G. Teply, and N. Whitehorn. Making maps of cosmic microwave background polarization for B-mode studies: the POLARBEAR example. *A&A*, 600:A60, Apr. 2017.

[15] G. Puglisi, D. Poletti, G. Fabbian, C. Baccigalupi, L. Heltai, and R. Stompor. Iterative map-making with two-level preconditioning for polarized Cosmic Microwave Background data sets. *ArXiv e-prints*, Jan. 2018.

[16] D. Ramanah, G. Lavaux, and B. Wandelt. Wiener filter reloaded: fast signal reconstruction without preconditioning. *Mon.Not.Roy.Astron.Soc.*, 468(2):1782–1793, 2017.

[17] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, PA, second edition, 2003.

[18] D. S. Seljebotn, T. Bærland, H. K. Eriksen, K.-A. Mardal, and I. K. Wehus. Multi-resolution Bayesian CMB component separation through Wiener-filtering with a pseudo-inverse preconditioner. *ArXiv e-prints*, Oct. 2017.

[19] K. M. Smith, O. Zahn, and O. Doré. Detection of gravitational lensing in the cosmic microwave background. *Phys. Rev. D*, 76(4):043510, Aug. 2007.

[20] Z. Strakoš and P. Tichý. Error estimation in preconditioned conjugate gradients. *BIT*, 45(4):789–817, 2005.

[21] M. Szydlarski, L. Grigori, and R. Stompor. Accelerating the cosmic microwave background map-making procedure through preconditioning. *A&A*, 572:A39, Dec. 2014.