We appreciate all the reviewers for their constructive comments.

R1

**Paper Weaknesses**

1 all the experimental results use different noise levels per color channels. It is interesting to see a direct comparison with the other method for the uniform case, when the same noise level affects all color channels

2 WNNM is very slow, a time complexity discussion and a runtime comparison with the other presented methods is a must;

3 how would the authors describe the practical importance of the proposed method in relation with the other compared methods?

4 I would appreciate comparison results with the recent DnCNN method from:

Zhang et al., "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising", 2016

5 line 535 : what means "tune for the best performance"? on which material the MC-WNNM performance was tuned?

6 for the real case visual results the authors could add the estimated noise level per each channel

1. Uniform noise levels for all channels

We added additive white Gaussian noise with standard deviation (std) of 25 to the R, G, B channels of the 24 color images from the Kodak PhotoCD dataset, and performed denoising with the methods listed in Table 1 of the main paper. The average results on PSNR of the compared methods are 24.08dB for CBM3D, 30.06 dB for MLP, 30.08 dB for TNRD, 25.82dB for NI, 27.15dB for NC, 29.91dB for WNNM-1, 30.02dB for WNNM-2, 30.32dB for WNNM-3 and MC-WNNM (the two methods are identical when the stds of R, G, B channel are equal).

1. Speed

We have added the complexity analysis and comparison on speed to the revised paper. Since the proposed MC-WNNM only need 2 iterations for denoising real color images, it takes about 200 seconds to process a real color image of size 512\*512\*3 and the speed is faster than WNNM (700 seconds), comparable to DnCNN (180 seconds), while slower than the other competing methods (e.g., CBM3D, MLP, TNRD, NI, NC).

1. Practical importance

The proposed method firstly incorporate the noise variance in different channels of color images and achieves better denoising performance on synthetic and real color image denoising tasks. It bases on and extends the original WNNM model to process color images and achieves better performance than state-of-the-art denoising methods on real color images.

1. Compare with DnCNN

We have performed the DnCNN method to the 24 color images in Kodak PhotoCD dataset and the 15 cropped images in [18]. The DnCNN can automatically deal with color images without knowing the noise standard deviation (std). We did not modify its settings or tune the parameters. On Kodak dataset, we add to the R, G, B channels of each color image the AWGN noise with stds of 40, 20, and 30, respectively. DnCNN achieves average 20.58dB at PSNR on the 24 images. If we perform DnCNN with known stds to process separately each channel of the color images, it achieves average 29.13dB at PSNR on the 24 images. The performance of DnCNN is inferior to that of the proposed MC-WNNM (29.31dB). On the 15 real noisy images in [18], DnCNN achieves averagely 33.86dB at PSNR, which is inferior to the proposed MC-WNNM method (37.71dB). On visual quality, DnCNN would either remain noise or generate artifacts while the proposed MC-WNNM remove the noise while maintaining the image details.

1. Tune parameters

We tune the tunable parameters to achieve highest PSNR results. We tune the number of iterations, the size of image patches, and the parameters in WNNM to achieve highest PSNR results for the methods of WNNM-1, WNNM-2, and WNNM-3. Besides, we also tune the parameters \rho and \mu for WNNM-3 and MC-WNNM.

1. Add noise level per channel

We had added the estimated noise level of each channel to the real color images.

R2

**Paper Weaknesses**

1 I have a general concern about the concept of nuclear norm in the general scenario. Please example the motivation of the use of nuclear norm. It is general to different noise types.

2 I am not sure that Eq.5 holds. The hypothesis under which it will work should be discussed and clarified.

3 I am not convinced that Eq. 6 has a such assumption. I hope author(s) can explain this design more in rebuttal.

1. The motivation of using nuclear norm

The low rank prior is widely used in many problems such as system identification, matrix completion, image processing (denoising, compression, background substraction, and inpainting, etc.). In image processing, the key motivation of using low rank prior is that, the rank of the matrix consisted of similar image patches is inherently lower than the number of patches. Since the rank of a matrix is non-convex, a common trick in optimization is to employ the convex envelope of matrix rank, i.e., the nuclear norm relaxation.

1. About Eq. 5

The maximum a-posteriori (MAP) estimation, i.e., Eq. 5, is an important research branch of Bayesian statistics and the fundamental of solutions to many inverse problems such as image restoration (e.g., denoising, super-resolution, deblurring, etc.). The MAP framework holds true and will work in general image restoration tasks.

1. About Eq. 6

As we have mentioned in the paper, we assume that the noise is independent among RGB channels and independently and identically distributed (i.i.d.) in each channel with Gaussian distribution. Hence, we have P(Y|X)=\prod\_{c=r,g,b} P(Y\_{c}|X\_{c}) and Eq. 6 holds true.

R3

**Paper Weaknesses**

1 well, perhaps the first sentence of the abstract could be rephrased - there exist many color denoising algo's - but they are difficult and not that well-known to the public.

2 and a statement that the code will be publicly available at xxx would be appreciated.

1. Abstract and code

Thank you very much for the encouraging comments. We had rephrased the abstract and added to the paper the publicly available address of the code in a revised version.